Contemporary Concepts in Physics
Volume 1

T. D. Lee

Particle Physics and Introduction to Field Theory

粒子物理和场论简引

harwood academic publishers

Particle Physics and Introduction to Field Theory

Contemporary Concepts in Physics

Volume 1
Particle Physics and Introduction to Field Theory

Additional volumes
in preparation.

ISSN: 0272-2488

Science Press (Beijing)

Particle Physics and Introduction to Field Theory

粒子物理和场论简引

T.D. Lee

Columbia University

harwood academic publishers
chur london new york

Published under license and distributed by:
Harwood Academic Publishers GmbH
Poststrasse 22
CH-7000 Chur, Switzerland

Editorial Office for the United Kingdom:
61 Grays Inn Road
London WC1X 8TL

Editorial Office for the United States of America:
Post Office Box 786
Cooper Station
New York, New York 10276

First Published: August 1981
Second Printing with Corrections: November 1982

A Chinese edition was published by
Science Press, Beijing

Library of Congress Cataloging in Publication Data
Lee, T.D., 1926—
Particle physics and introduction to field theory.
(Contemporary concepts in physics; v. 1)
1. Particles (Nuclear physics) 2. Field theory (Physics)
3. Symmetry (Physics) 4. Nuclear reactions. I. Title. II. Series.
QC793.2.L43 1981 539.7'21 80-27348
ISBN 3-7186-0032-3 AACR1
ISBN3-7186-0033-1 (pbk.)

Library of Congress catalog card number 80-27348. ISSN 0272-2488. ISBN 3-7186 0032-3 (cloth); 3-7186-0033-1 (paperback).

Printed in the United States of America.

TO JEANNETTE

Contents

PREFACE TO THE SERIES

The series of volumes, Concepts in Contemporary Physics, of which the present book by T. D. Lee is the first, is addressed to the professional physicist and to the serious graduate student of physics. The subjects to be covered will include those at the forefront of current research. It is anticipated that the various volumes in the series will be rigorous and complete in their treatment, supplying the intellectual tools necessary for the appreciation of the present status of the areas under consideration and providing the framework upon which future developments may be based. An examination of Professor Lee's book reveals that these criteria are more than amply fulfilled and that his lucid, unique and profound presentation of the entire subject of elementary particle physics and quantum field theory is an excellent model for the volumes to follow in this series.

H. Primakoff

PREFACE

In the spring of 1979 I was invited by Academia Sinica to give lectures on particle physics and statistical mechanics in Beijing. Parts of the lecture notes have been published in Chinese, and the particle physics section evolved into this volume.

The intention of the course was to bring both theoretical and experimental students of physics to the forefront of this very exciting and active field. Because of the different backgrounds of those who attended my lectures, this book is self-contained. Whenever possible, I have adopted an approach that is more pragmatic than axiomatic. All derivations are done explicitly, which at times to a more sophisticated reader may appear pedantic.

Among the topics not treated are the details of renormalization in field theory, the use of dispersion techniques in particle physics and the very beautiful topological soliton solutions in gauge theories. Fortunately, some of these subjects are well covered by the well-known books of Bjorken and Drell and the recent volume by Itzykson and Zuber.

Many people have given me valuable suggestions, and I wish to thank in particular N. H. Christ and A. H. Mueller. It would not have been possible for me to complete this volume without the essential assistance of Irene Tramm.

T. D. Lee

New York
July 1981

Chapter 1

MECHANICS OF A FINITE SYSTEM (REVIEW)

1.1 Classical Mechanics

Let us first consider a classical system of particles whose generalized coordinates are q_i $(i = 1,2, \cdots, N)$. For example, $N = 3n$ if we have n particles in three dimensions. Suppose the Lagrangian is

$$L = L(q_i , \dot{q}_i) \tag{1.1}$$

where $\dot{q}_i$ denotes the time derivative of q_i . The Lagrangian equation of motion is given by the variational principle

$$\delta \int_{t_1}^{t_2} L dt = 0 \ , \tag{1.2}$$

in which δ denotes the variation with the boundary condition $\delta q_i = 0$ at the initial and final times, t_1 and t_2 . This is the well-known action principle, and it leads to Lagrange's form of the equations of motion

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \ . \tag{1.3}$$

The generalized momentum p_i is

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i} \ . \tag{1.4}$$

The Hamiltonian of the system is then given by

$$H(q_i , p_i) \equiv \sum_i p_i \dot{q}_i - L \ . \tag{1.5}$$

The transformation relating $L(q_i, \dot{q}_i)$ and $H(q_i, p_i)$ is called the Legendre transformation, in which L is regarded as a function of q_i and $\dot{q}_i$, while H is a function of q_i and p_i.

Quite often, we shall adopt the convention that repeated indices are supposed to be summed over. Thus, (1.5) can be simply written as

$$H(q_i, p_i) = p_i \dot{q}_i - L \quad .$$

From (1.3)-(1.5) one readily obtains Hamilton's form of the equations of motion

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \quad , \quad \dot{q}_i = \frac{\partial H}{\partial p_i} \quad . \tag{1.6}$$

1.2 Quantization

Next, we shall discuss the quantization of the system. Assuming that the Hamiltonian $H(q_i, p_i)$ is given for a classical system, in order to quantize this system we first regard $q_i(t)$ and $p_i(t)$ as operators which satisfy the commutation relations

$$\begin{aligned} &[q_i(t), p_j(t)] = i\delta_{ij} \\ \text{and} \qquad &[q_i(t), q_j(t)] = [p_i(t), p_j(t)] = 0 \end{aligned} \tag{1.7}$$

in which $[A, B] = AB - BA$ and δ_{ij} is the Kronecker symbol

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad .$$

In passing from classical to quantum mechanics, each physical observable becomes a Hermitian operator. If we represent a Hermitian operator in its matrix form, then its matrix elements satisfy

$$A_{ij} = (A^\dagger)_{ij} \equiv (A^*)_{ji} = (A_{ji})^*$$

where † denotes Hermitian conjugation and * complex conjugation.

Thus we have

$$q_i = q_i^\dagger \ , \quad p_i = p_i^\dagger \ , \quad L = L^\dagger \quad \text{and} \quad H = H^\dagger \ . \tag{1.8}$$

In classical mechanics the time dependence of q_i and p_j is given by Hamilton's equations (1.6). In quantum mechanics the time derivative $\dot{O}$ of any operator $O(t)$ is determined by Heisenberg's equation

$$[H, O(t)] = -i\dot{O}(t) \ . \tag{1.9}$$

By regarding H as a polynomial of q_i and p_i, one can verify that Heisenberg's equation leads to the same Hamilton equations when $O(t) = q_i$ and p_i. [See the example below.]

In classical mechanics q_i and p_j commute. Thus there can be ambiguities in passing from the classical Hamiltonian $H(q_i, p_j)$ to its quantum-mechanical form. For example

$$H_1 = p^3q^2 + q^2p^3 \qquad \text{and} \qquad H_2 = 2pqpqp$$

represent identical systems in classical mechanics, but in quantum mechanics they correspond to different Hamiltonians. One may therefore ask which form one should choose. The answer is that these are two different quantum-mechanical systems, each having the same classical limit. Knowing the classical limit does not always imply a unique determination of the corresponding quantum-mechanical system. For a realistic physical system, only through direct comparison between the experimental result and the theoretical analysis can one be sure which Hamiltonian form is the correct one.

Example. The harmonic oscillator

The simplest harmonic oscillator is one with unit frequency. Its Lagrangian is

$$L = L(q, \dot{q}) = \tfrac{1}{2}(\dot{q}^2 - q^2) \ .$$

Hence

$$p = \frac{\partial L}{\partial \dot{q}} = \dot{q} \ , \tag{1.10}$$

and therefore

$$H(q, p) = \tfrac{1}{2}(p^2 + q^2) \ .$$

By using Hamilton's equations (1.6), one obtains

$$\dot{q} = \frac{\partial H}{\partial p} = p \quad \text{and} \quad \dot{p} = -\frac{\partial H}{\partial q} = -q \ . \tag{1.11}$$

In classical mechanics, the commutators between these functions are all zero. They are called c. number (for commuting) functions.

To quantize the system we change all the above c. number functions to the appropriate q. number (for quantum) operators. By using Heisenberg's equation (1.9) we derive

$$-i\dot{p} = [H, p] = \tfrac{1}{2}[q^2, p]$$
$$= \tfrac{1}{2}(q(qp - pq) - (pq - qp)q) = iq \ ,$$

i.e.,

$$\dot{p} = -q \ .$$

Likewise, we find

$$-i\dot{q} = [H, q] = -ip \ ,$$

i.e.,

$$\dot{q} = p \ .$$

Thus, Heisenberg's equation gives the identical result as Hamilton's equations.

To analyze the eigenvalue problem we introduce

$$a \equiv \frac{1}{\sqrt{2}}(q + ip) \ . \tag{1.12}$$

In accordance with (1.8), we have $q = q^\dagger$ and $p = p^\dagger$, and therefore the Hermitian conjugate of a is

$$a^\dagger = \frac{1}{\sqrt{2}}(q - ip) \ . \tag{1.13}$$

We may solve q and p in terms of a and $a^\dagger$:

$$q = \frac{1}{\sqrt{2}} (a + a^\dagger) \quad , \tag{1.14}$$

$$p = \frac{-i}{\sqrt{2}} (a - a^\dagger) \quad . \tag{1.15}$$

Because $[p, q] = -i$, we find

$$[a, a^\dagger] = 1 \quad . \tag{1.16}$$

Moreover,

$$a^\dagger a = \tfrac{1}{2}(q - ip)(q + ip)$$
$$= \tfrac{1}{2}\{q^2 + p^2 - i(pq - qp)\} = H - \tfrac{1}{2} \quad ,$$

which leads to $H = a^\dagger a + \frac{1}{2}$. Let

$$N \equiv a^\dagger a \quad , \tag{1.17}$$

then the Hamiltonian H can also be written as

$$H = N + \tfrac{1}{2} \quad . \tag{1.18}$$

The operator N is non-negative since its expectation value over any state vector is an absolute value squared and is therefore $\geqslant 0$.

We shall now show that the eigenvector $| n >$ of H satisfies

$$H | n > = (n + \tfrac{1}{2}) | n > \tag{1.19}$$

in which n can be any positive integer $0, 1, 2, \cdots$. Furthermore, let $| 0 >$ be the eigenvector with the smallest eigenvalue, i.e.

$$H | 0 > = \tfrac{1}{2} | 0 > \quad ; \tag{1.20}$$

then the other eigenvectors can be written as

$$| n > = \frac{1}{\sqrt{n!}} (a^\dagger)^n | 0 > \quad . \tag{1.21}$$

Proof. Let $|\ >$ be any eigenvector of N :

$$N | \ > = \ell | \ > \tag{1.22}$$

where ℓ is a number. Because

$$Na^\dagger = a^\dagger a a^\dagger = a^\dagger (a^\dagger a + 1) = a^\dagger (N + 1) \tag{1.23}$$

and

$$Na = a^\dagger aa = (aa^\dagger - 1)a = a(N-1) \quad , \tag{1.24}$$

it follows that $Na^\dagger|> = (\ell+1)\,a^\dagger|>$ and $Na|> = (\ell-1)\,a|>$. Next, replace $|>$ respectively by $a^\dagger|>$ in the first equation and by $a|>$ in the second. After repeating this process n times, we find

$$N(a^\dagger)^n|> = (\ell+n)\,(a^\dagger)^n|> \tag{1.25}$$

and

$$Na^n|> = (\ell-n)\,a^n|> \quad . \tag{1.26}$$

Now, if $\ell \neq$ integer, by choosing in (1.26) n = integer $> \ell$ we would obtain a negative eigenvalue for N ; that is impossible since, as noted before, N is non-negative. Hence ℓ = integer, and according to (1.26)

$$|0> \equiv a^\ell|>$$

satisfies

$$N|0> = 0 \quad . \tag{1.27}$$

Furthermore, zero must be the smallest eigenvalue of N . By setting $|0>$ in place of $|>$ in (1.25), we find

$$N(a^\dagger)^n|0> = n(a^\dagger)^n|0>$$

where n can be any positive integer. Thus, (1.19) - (1.21) are proved. Let us choose the normalization of the state $|0>$ so that $<0|0> = 1$. Then from (1.21) all other states $|n>$ are also normalized: $<n|n> = 1$.

From (1.24) and (1.27), we see that

$$Na|0> = -a|0> \quad .$$

Since N is non-negative, we must have

$$a|0> = 0 \quad . \tag{1.28}$$

In the coordinate representation p is $-i\frac{\partial}{\partial q}$. Let $\psi(q) = <q|0>$. Equation (1.28) becomes

$$a\psi = \frac{1}{\sqrt{2}}(q + ip)\psi = \frac{1}{\sqrt{2}}(q + \frac{\partial}{\partial q})\psi = 0$$

which determines the solution $\psi(q) \propto e^{-\frac{1}{2}q^2}$. Therefore, not only does $|0>$ exist, it is also non-degenerate. This completes the proof.

According to (1.21)

$$a^\dagger |n> = \sqrt{n+1}\ |n+1>\ , \tag{1.29}$$

therefore $a^\dagger |n-1> = \sqrt{n}\ |n>$, which leads to

$$a|n> = \frac{aa^\dagger}{\sqrt{n}}|n-1> = \frac{1}{\sqrt{n}}(a^\dagger a + 1)|n-1>\ ,$$

i.e.,

$$a|n> = \sqrt{n}\ |n-1>\ . \tag{1.30}$$

The operator $N = a^\dagger a$ is commonly called the occupation-number operator. Because of (1.29) and (1.30), $a^\dagger$ is called the creation operator and a the annihilation operator.

In this example of the harmonic oscillator we may consider a space whose basis vectors are the eigenvectors $|0>, |1>, \cdots$. Due to the orthonormality of these vectors we may write

$$|0> = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \end{pmatrix}\ , \qquad |1> = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix}\ , \cdots\ .$$

The corresponding matrix forms of N, $a^\dagger$ and a are

$$N = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 2 & \cdots \\ \vdots & \vdots & \vdots & \vdots\vdots \end{pmatrix},$$

$$a^\dagger = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ \sqrt{1} & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots\vdots \end{pmatrix}, \quad a = \begin{pmatrix} 0 & \sqrt{1} & 0 & \cdots \\ 0 & 0 & \sqrt{2} & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots\vdots \end{pmatrix}.$$

(1.31)

Exercise. Repeat the same steps for a harmonic oscillator whose Hamiltonian is $H = \frac{1}{2m} p^2 + \frac{1}{2} k q^2$ where m and k are constants.

1.3 Some General Theorems

Let H be a Hermitian operator in a linear space

$$\{v_a\} \quad , \qquad a = 0,1,2, \cdots \tag{1.32}$$

where the basis vectors v_a form an orthonormal set. A vector $|\ >$ in this space is sometimes called a state vector, or simply a state. We shall assume that the Hermitian operator H is bounded from below; i.e., for any state $|\ >$ the ratio $\frac{<|H|>}{<|>}$ is always larger than a fixed constant c. Its eigenvector equation can be written as $H\,|\,a> = E_a\,|\,a>$, and its eigenvalues E_a can be arranged so that (for $a = 0,1,2, \cdots$)

$$E_0 \leqslant E_1 \leqslant \cdots \leqslant E_m \leqslant E_{m+1} \leqslant \cdots \ . \tag{1.33}$$

Since H is Hermitian, we can always choose

$$<a\,|\,a'> = \delta_{aa'} \tag{1.34}$$

for any eigenvectors $|\,a>$ and $|\,a'>$ of H.

Theorem 1. The minimum of $\frac{<|H|>}{<|>}$ is

(i) E_0, if $|\ >$ can be any state vector,

(ii) E_1, if $|\ >$ can be any state vector that satisfies the constraint $<0\,|\,> = 0$,

(iii) E_n, if $|\ >$ can be any state vector that satisfies the constraints $<0\,|\,> = <1\,|\,> = \cdots = <n-1\,|\,> = 0$.

Proof. Let

$$E \equiv \frac{<|H|>}{<|>} \ ,$$

and denote

$$\psi = |\ > \ , \quad \psi^\dagger = <|\ .$$

To find the minimum of E we may consider the variation $\psi \to \psi + \delta\psi$. The corresponding variation in E is

$$\delta E = \frac{1}{\psi^\dagger\psi}(\delta\psi^\dagger H\psi + \psi^\dagger H\delta\psi) - \frac{\psi^\dagger H\psi}{\psi^\dagger\psi}\left(\frac{\delta\psi^\dagger\psi}{\psi^\dagger\psi} + \frac{\psi^\dagger\delta\psi}{\psi^\dagger\psi}\right)$$
$$= \frac{1}{\psi^\dagger\psi}[\delta\psi^\dagger(H-E)\psi + \psi^\dagger(H-E)\delta\psi] \ .$$

Set $f = (H-E)\psi$. If $f = 0$, then clearly $\delta E = 0$. If $f \neq 0$, we may choose $\delta\psi = \epsilon f$ where ϵ is an infinitesimal real quantity. Thus, the above equation can be written as

$$\delta E = \frac{2}{\psi^\dagger\psi}\epsilon f^\dagger f \ .$$

The necessary condition for E to be a minimum is $\delta E = 0$ for arbitrary $\delta\psi$. Hence $f = 0$, which implies $(H-E)\psi = 0$. Since E_0 is the smallest eigenvalue of H, (i) is then established.

In (1.32) we may choose $v_0 = |0>$. Let us consider the subspace $\{v_i\}$ which is spanned by all v_i with $i \geqslant 1$. By definition, all vectors orthogonal to $|0>$ are in this subspace. Furthermore, for any v_i $(i \geqslant 1)$, H satisfies

$$<v_0|H|v_i> = E_0<v_0|v_i> = 0 \ .$$

Hence Hv_i also belongs to this subspace $\{v_i\}$. Now consider the Hamiltonian in this subspace. By following the same argument as that used in proving (i) we can establish (ii), and likewise also (iii).

We call a set of basis vectors $\{|a>\}$ complete if for any state vector $|>$, there exists a set of numerical constants $\{C_a\}$ such that

$$\lim_{m\to\infty} <R_m|R_m> = 0 \ ,$$

where

$$|R_m> \equiv |> - \sum_{a=0}^{m} C_a|a> \ .$$

Theorem 2. If a Hermitian operator H is bounded from below, but not from above (i.e., for any real constant c there exists a state vector $|\ >$ such that $\frac{<|\ H\ |>}{<|\ |>}$ is larger than c), then the set of all its eigenvectors $\{|\ a>\}$ is complete.

Proof. Since H is a Hermitian operator, we can choose its eigenvectors to satisfy (1.34). The theorem is obvious if $\{|\ a>\}$ spans a finite dimensional space. We need only consider the case when $\{|\ a>\}$ spans a space of infinite dimensions.

Let us arrange the eigenvalues of H in the form (1.33). Because H is bounded from below, by replacing $H \to H +$ a constant we can set

$$E_0 > 0 \ .$$

Let us choose $C_a = <a\ |\ >$. Consequently

$$|\ R_m> = |\ > - \sum_{a=0}^{m} C_a\ |\ a>$$

satisfies

$$<a\ |\ R_m> = 0 \qquad \text{when} \qquad a \leqslant m \ .$$

From Theorem 1 we have

$$\frac{<R_m\ |\ H\ |\ R_m>}{<R_m\ |\ R_m>} \geqslant E_{m+1} \geqslant E_m \ , \tag{1.35}$$

Because H is not bounded from above, it follows that

$$E_m \to \infty \qquad \text{when} \qquad m \to \infty \ . \tag{1.36}$$

In addition,

$$<R_m\ |\ H\ |\ R_m> = (<|\ - \sum_a C_a^* <a|\ )\ H\ (|\ > - \sum_b C_b\ |\ b>)$$

$$= <|\ H\ |\ > - \sum_a C_a^* <a|\ H\ |\ > - \sum_b C_b <|\ H\ |b> + \sum_{a,b} C_a^* C_b <a|\ H\ |b>$$

where the sum extends over a and $b = 1, 2, \cdots, m$. Since $|\ a>$

and $|\ b\ >$ are eigenstates of H, this leads to

$$<R_m\ |\ H\ |\ R_m> = <\ |\ H\ |\ > - \sum_a C_a^* C_a E_a - \sum_b C_b C_b^* E_b + \sum_a C_a^* C_a E_a$$

$$= <\ |\ H\ |\ > - \sum_b C_b C_b^* E_b$$

which is $\leqslant$ the first term $<\ |\ H\ |\ >$, since the second term $-\sum_b C_b C_b^* E_b$ is always negative. Thus, by using (1.35), we find

$$<R_m\ |\ R_m> \leqslant \frac{1}{E_m} <R_m\ |\ H\ |\ R_m> \leqslant \frac{1}{E_m} <\ |\ H\ |\ >. \quad (1.37)$$

Because $<\ |\ H\ |\ >$ is independent of m and $<R_m\ |\ R_m>$ is positive, by using (1.36) we obtain

$$<R_m\ |\ R_m> \to 0 \qquad \text{when} \qquad m \to \infty\ .$$

This establishes Theorem 2. Therefore the set of all eigenvectors of H can be used as the complete set of basis vectors in the Hilbert space.

The following are a few examples:

(i) $H = \frac{1}{2}p^2 + V(x)$, in which $V(x)$ has a lower bound.

In the x-representation $p^2 = -\frac{d^2}{dx^2}$. Choose

$$<x\ |\ > \propto e^{-x^2/\lambda^2}\ ,$$

then, when $\lambda \to 0$,

$$\frac{<\ |\ p^2\ |\ >}{<\ 1\ >} \to \infty\ .$$

Hence, H is bounded from below but not from above. The set of all its eigenvectors forms a complete set of functions.

(ii) Consider a circle in a two-dimensional space. Let $H = p_\phi^2 = -\frac{d^2}{d\phi^2}$ where ϕ is the angular variable which varies from 0 to 2π.

Let $|\ m\ >$ be an eigenstate of H. In the ϕ-representation, we may denote $\psi_m(\phi) = <\phi\ |\ m\ >$, which is a periodic function of ϕ. The solutions of the eigenstate equation, $H\psi_m = m^2 \psi_m$, are

$$\psi_m = \frac{1}{\sqrt{2\pi}} e^{im\phi} \qquad \text{where} \qquad m = 0, \pm 1, \cdots\ .$$

From Theorem 2 we know that, except for a set of points of zero measure, any function $f(\phi)$ can be expanded in terms of these eigenfunctions:

$$f(\phi) = \sum C_m e^{im\phi} \quad .$$

This is the main content of the well-known Fourier theorem.

(iii) Next, we consider the surface of a unit sphere in a three-dimensional space. Let $H = -\nabla^2$ on the surface; i.e., in terms of the spherical coordinates

$$H = -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) - \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \quad .$$

The eigenfunctions of H are called the spherical harmonics $Y_{\ell m}(\theta, \phi)$ which satisfy

$$H Y_{\ell m}(\theta, \phi) = \ell(\ell+1) Y_{\ell m}(\theta, \phi) \quad , \tag{1.38}$$

where $\ell = 0, 1, 2, \cdots$. The ϕ-dependence of $Y_{\ell m}$ is

$$Y_{\ell m} \propto e^{im\phi} \tag{1.39}$$

with $m = 0, \pm 1, \pm 2, \cdots \pm \ell$. Theorem 2 tells us that, except for a set of points of zero measure, any function $f(\theta, \phi)$ can be expanded in terms of $Y_{\ell m}$:

$$f(\theta, \phi) = \sum C_{\ell m} Y_{\ell m}(\theta, \phi) \quad .$$

Equations (1.38) and (1.39) determine $Y_{\ell m}$ up to a multiplicative factor. The usual choices are

$$Y_{0,0} = \frac{1}{\sqrt{4\pi}} \quad , \quad Y_{1,\pm 1} = \mp\sqrt{\frac{3}{8\pi}}\sin\theta\, e^{\pm i\phi} \quad ,$$

$$Y_{1,0} = \sqrt{\frac{3}{4\pi}}\cos\theta \quad , \quad \text{etc.} \tag{1.40}$$

The application of Theorem 2 is quite general. The space spanned by such a complete set of vectors is called Hilbert space.

Problem 1.1 Let H be a Hermitian operator that is bounded from below. Arrange its eigenvalues $E_0, E_1, \cdots$ and the corresponding eigenvectors $|0>, |1>, \cdots$ in the order given by (1.33).

(i) Let $|b>$ be an arbitrarily chosen vector. Define $F(b)$ to be the minimum of $\frac{<|H|>}{<|>}$ where $|>$ can be any vector that satisfies $<b|> = 0$. By varying $|b>$, prove that the maximum of $F(b)$ is E_1, the second lowest eigenvalue of H.

Hint: To find $F(b)$, consider the vector

$$<b|1>|0> - <b|0>|1> .$$

(ii) Let $|b_1>, |b_2>, \cdots, |b_n>$ be n arbitrarily chosen linearly independent vectors. Define $F(b_1, b_2, \cdots, b_n)$ to be the minimum of $\frac{<|H|>}{<|>}$ where $|>$ can be any vector that satisfies $<b_1|> = <b_2|> = \cdots = <b_n|> = 0$. Prove that the maximum of $F(b_1, b_2, \cdots, b_n)$ is E_n. (maximum-minimum principle)

Problem 1.2 (i) In the above problem, suppose a constraint C is imposed on all state vectors. Correspondingly, all eigenvalues and eigenvectors will be changed: $E_n \to E_n'$ and $|n> \to |n'>$. By applying the maximum-minimum principle, prove that $E_0 \leqslant E_0'$, $E_1 \leqslant E_1', \cdots, E_n \leqslant E_n', \cdots$.

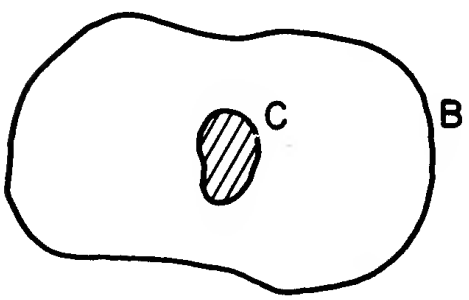


Fig. 1.1

(ii) Consider the vibration of a membrane with a fixed boundary B. The characteristic frequency ω_n is determined by $-\nabla^2\phi = \omega_n^2\phi$ where the vibrational amplitude ϕ is zero at the boundary. Arrange these frequencies in the order

$\omega_0 \leqslant \omega_1 \leqslant \omega_2 \leqslant \cdots$. Impose the constraint that ϕ is also zero inside a closed curve C within the membrane, as shown in Fig. 1.1 ; correspondingly, the characteristic frequency ω_n is changed to ω_n'. Show that $\omega_n \leqslant \omega_n'$ for all n .

References

Some standard textbooks on subjects discussed in this chapter are:

R. Courant and D. Hilbert, Methods of Mathematical Physics, 2 vols. (New York, Interscience Publishers, Inc., 1962).

P. A. M. Dirac, Quantum Mechanics (Oxford, The Clarendon Press, 1958).

E. T. Whittaker, Analytical Dynamics (Cambridge, The University Press, 1960).

Chapter 2

THE SPIN-0 FIELD

2.1 General Discussion

Now we turn to the quantization of a local field theory. Let $\phi(x)$ be a local field where $x = x_\mu = (\vec{r}, it)$; i.e., $x_i = \vec{r}_i$ for $i = 1,2,3$ and $x_4 = it$. If, under the Lorentz transformation ϕ is invariant, then we call it a spin-0 field. In addition, we call ϕ a scalar field if it does not change sign under the space inversion; otherwise, a pseudoscalar field. Let us begin our discussions by considering a Hermitian field:

$$\phi(x) = \phi(\vec{r}, t) = \phi^\dagger(x) \quad .$$

We may first enclose the whole system in a finite rectangular box of size Ω, and assume ϕ to satisfy the periodic boundary condition, and then in the end let $\Omega \to \infty$. It may be emphasized that this particular procedure is no less physical than that in which Ω is set to be infinite at the beginning. For all we know our universe may well be finite, since within the present experimental degree of accuracy in particle physics it is not even remotely possible to determine its size, let alone its boundary conditions. Therefore, different routes allowing Ω to go to infinity should lead to the same theoretical result.

Let the Lagrangian density of the system be

$$\mathcal{L} = -\tfrac{1}{2}\left(\frac{\partial\phi}{\partial x_\mu}\right)^2 - V(\phi) \quad , \tag{2.1}$$

in which, here as well as in later discussions, the repeated index μ is summed over from 1 to 4. Therefore, in the above expression,

$$\left(\frac{\partial\phi}{\partial x_\mu}\right)^2 = \frac{\partial\phi}{\partial x_\mu}\,\frac{\partial\phi}{\partial x_\mu} = (\vec{\nabla}\phi)^2 - \dot{\phi}^2 \quad .$$

The Lagrangian L is given by

$$L = \int_\Omega \mathcal{L}\, d^3r = \int_\Omega \mathcal{L}\,(\phi(\vec{r}, t),\ \dot{\phi}(\vec{r}, t))\, d^3r \quad , \tag{2.2}$$

which is regarded as a functional $L(\phi, \dot{\phi})$ of $\phi(\vec{r}, t)$ and its time derivative $\dot{\phi}(\vec{r}, t)$. By comparing it with (1.1), we see that ϕ corresponds to the generalized coordinate with $\dot{\phi}$ the corresponding velocity. The main difference is that while in (1.1) the index i is discrete and of finite value, in the case of a field the corresponding index is $\vec{r}$, which is continuous and has an infinite number of values.

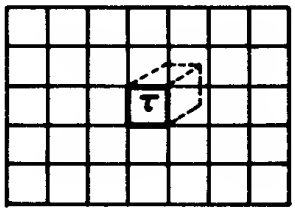


Fig. 2.1 Division of Ω into N tiny cubes, each of size τ.

For convenience, we may divide Ω into many small cubes of size τ, as in Fig. 2.1. Let the value of $\phi(\vec{r}, t)$ in each particular small cube τ be represented by $\phi(\vec{r}_i, t) \equiv \phi_i(t)$ where $\vec{r}_i$ is the coordinate of any arbitrarily chosen fixed point in this little cube. Set $q_i(t) \equiv \tau\,\phi_i(t)$. Equation (2.2) will now be written as

$$L = \tfrac{1}{2}\int_\Omega \dot\phi^2\, d^3r - \int_\Omega [\tfrac{1}{2}(\nabla\phi)^2 + V(\phi)]\, d^3r$$
$$= \tfrac{1}{2}\sum_i \frac{\dot q_i^{\,2}}{\tau} - \cdots ,$$

where the $\cdots$ denotes terms independent of $\dot q_i$. Therefore the corresponding generalized momentum is

$$p_i(t) = \frac{\partial L}{\partial \dot q_i} = \frac{\dot q_i}{\tau} = \dot\phi_i(t) \equiv \Pi(\vec r_i, t) \quad . \tag{2.3}$$

When $\tau \to 0$, by using (1.5) we find the Hamiltonian to be

$$H = \sum_i p_i \dot q_i - L = \int [\tfrac{1}{2}\Pi^2 + \tfrac{1}{2}(\nabla\phi)^2 + V(\phi)]\, d^3r . \tag{2.4}$$

According to the general rules of quantization

$$[p_i(t),\ q_j(t)] = -i\delta_{ij} ,$$

we have

$$[\Pi(\vec r_i, t),\ \phi(\vec r_j, t)] = -i\frac{\delta_{ij}}{\tau} ,$$

which leads to, when $\tau \to 0$,

$$[\Pi(\vec r, t),\ \phi(\vec r', t)] = -i\delta^3(\vec r - \vec r') , \tag{2.5}$$

where $\delta^3(\vec r - \vec r')$ is the three-dimensional Dirac δ-function. The definition of the δ-function is

$$\delta^3(\vec r - \vec r') = 0 \quad \text{if} \quad \vec r \neq \vec r'$$

and

$$\int \delta^3(\vec r - \vec r')\, d^3r = 1$$

in which the integration extends over any volume that includes the point $\vec r'$. In (2.5) both points $\vec r$ and $\vec r'$ are assumed to be inside the box Ω. Likewise, because $[q_i(t),\ q_j(t)] = [p_i(t),\ p_j(t)] = 0$,

$$[\phi(\vec r, t),\ \phi(\vec r', t)] = [\Pi(\vec r, t),\ \Pi(\vec r', t)] = 0 . \tag{2.6}$$

The equation of motion of the field remains given by the same Heisenberg equation, (1.9). By setting $O(t)$ to be $\phi(\vec r, t)$ in (1.9) we find

$$[H, \phi(\vec r, t)] = -i\dot\phi(\vec r, t) .$$

On account of (2.4)-(2.6) we see that the lefthand side is $-i\Pi(\vec{r}, t)$, and that leads to

$$\Pi(\vec{r}, t) = \dot{\phi}(\vec{r}, t) \quad , \tag{2.7}$$

in agreement with (2.3). Likewise, by setting $O(t) = \Pi(\vec{r}, t)$ in Heisenberg's equation (1.9) we have

$$[H, \Pi(\vec{r}, t)] = -i\dot{\Pi}(\vec{r}, t) \quad . \tag{2.8}$$

Now, (2.5) can also be written as

$$[\phi(\vec{r}, t), \Pi(\vec{r}', t)] = i\delta^3(\vec{r} - \vec{r}') = i\frac{\delta\phi(\vec{r}, t)}{\delta\phi(\vec{r}', t)} \quad ,$$

where, in the variational derivative of the last expression, we may regard

$$\phi(\vec{r}, t) = \int \delta^3(\vec{r} - \vec{r}')\,\phi(\vec{r}', t)\,d^3r \quad .$$

Consequently,

$$[H, \Pi(\vec{r}, t)] = i\frac{\delta H}{\delta\phi(\vec{r}, t)} = i(-\nabla^2\phi + \frac{dV}{d\phi}) \quad ,$$

which, together with (2.7) and (2.8), gives

$$\ddot{\phi} - \nabla^2\phi + \frac{dV}{d\phi} = 0 \quad . \tag{2.9}$$

The same equation of motion can also be derived by using the variational principle

$$\delta \int L dt = \delta \int \mathcal{L}\, d^4x = 0 \quad . \tag{2.10}$$

Of course, this is not an accident. The underlying reason is the same as that discussed in Chapter 1 for a finite system.

Thus we see that the quantum field theory is merely an extension of the ordinary quantum mechanics of a finite system to an infinite system.

2.2 Fourier Expansion (Free or Interacting Fields)

At any given time t the operators $\phi(\vec{r}, t)$ and $\Pi(\vec{r}, t)$ can be expanded in terms of the Fourier series:

$$\phi(\vec{r}, t) = \sum_{\vec{k}} \frac{1}{\sqrt{\Omega}} e^{i\vec{k}\cdot\vec{r}} q_{\vec{k}}(t) \tag{2.11}$$

and

$$\Pi(\vec{r}, t) = \sum_{\vec{k}} \frac{1}{\sqrt{\Omega}} e^{i\vec{k}\cdot\vec{r}} p_{-\vec{k}}(t) \tag{2.12}$$

in which $q_{\vec{k}}(t)$ and $p_{\vec{k}}(t)$ are time-dependent operators in the Hilbert space, and components of $\vec{k}$ are given by

$$k_i = \frac{2\pi \ell_i}{L_i} , \qquad i = 1, 2, 3 \tag{2.13}$$

with

$$\ell_i = 0, \pm 1, \pm 2, \cdots \tag{2.14}$$

and L_1, L_2, L_3 are, respectively, the lengths of the three sides of the rectangular box Ω. The validity of this expansion depends only on the completeness property of the Fourier series, discussed in the last chapter. [We may view the operators $\phi(\vec{r}, t)$ and $\Pi(\vec{r}, t)$ as matrices; each of their matrix elements is a c. number function of $\vec{r}$ and t which at any given time t can be expanded in terms of the Fourier series, and that results in the above expansions.] Since ϕ and Π are Hermitian operators, i.e., $\phi(\vec{r}, t) = \phi^\dagger(\vec{r}, t)$ and $\Pi(\vec{r}, t) = \Pi^\dagger(\vec{r}, t)$, we have

$$q_{\vec{k}}(t) = q^\dagger_{-\vec{k}}(t) , \qquad p_{\vec{k}}(t) = p^\dagger_{-\vec{k}}(t) . \tag{2.15}$$

Let us define

$$a_{\vec{k}}(t) \equiv \sqrt{\frac{\omega}{2}} \left(q_{\vec{k}} + \frac{i}{\omega} p_{-\vec{k}}\right) , \tag{2.16}$$

where

$$\omega \equiv \sqrt{\vec{k}^2 + m^2} \tag{2.17}$$

and m is an arbitrarily chosen real parameter. Because of (2.15), the Hermitian conjugate of $a_{\vec{k}}(t)$ is

$$a^\dagger_{\vec{k}}(t) = \sqrt{\frac{\omega}{2}} \left(q_{-\vec{k}} - \frac{i}{\omega} p_{\vec{k}}\right) . \tag{2.18}$$

By changing the sign of $\vec{k}$ we obtain

$$a^\dagger_{-\vec{k}}(t) = \sqrt{\frac{\omega}{2}} \left(q_{\vec{k}} - \frac{i}{\omega} p_{-\vec{k}}\right) . \tag{2.19}$$

From (2.16) and (2.19) we may express $q_{\vec{k}}$ and $p_{-\vec{k}}$ in terms of $a_{\vec{k}}$ and $a^{\dagger}_{-\vec{k}}$:

$$q_{\vec{k}}(t) = \frac{1}{\sqrt{2\omega}} [a_{\vec{k}}(t) + a^{\dagger}_{-\vec{k}}(t)]$$

and

$$p_{-\vec{k}}(t) = \frac{-i\omega}{\sqrt{2\omega}} [a_{\vec{k}}(t) - a^{\dagger}_{-\vec{k}}(t)] \quad .$$

By substituting these expressions into (2.11) and (2.12) we find

$$\phi(\vec{r}, t) = \sum \frac{1}{\sqrt{2\omega\Omega}} [a_{\vec{k}}(t)\, e^{i\vec{k}\cdot\vec{r}} + a^{\dagger}_{\vec{k}}(t)\, e^{-i\vec{k}\cdot\vec{r}}] \qquad (2.20)$$

and

$$\Pi(\vec{r}, t) = \sum \frac{-i\omega}{\sqrt{2\omega\Omega}} [a_{\vec{k}}(t)\, e^{i\vec{k}\cdot\vec{r}} - a^{\dagger}_{\vec{k}}(t)\, e^{-i\vec{k}\cdot\vec{r}}] \quad . \qquad (2.21)$$

Next, we observe that on account of (2.13) and (2.14),

$$\Omega^{-1} \int e^{i(\vec{k}-\vec{k}')\cdot\vec{r}}\, d^3r = \delta_{\vec{k},\vec{k}'} \qquad (2.22)$$

which is 0 if $\vec{k} \neq \vec{k}'$ and is 1 if $\vec{k} = \vec{k}'$. Therefore, from (2.11) and (2.12) we have

$$q_{\vec{k}}(t) = \frac{1}{\sqrt{\Omega}} \int e^{-i\vec{k}\cdot\vec{r}}\, \phi(\vec{r}, t)\, d^3r$$

and

$$p_{\vec{k}}(t) = \frac{1}{\sqrt{\Omega}} \int e^{i\vec{k}\cdot\vec{r}}\, \Pi(\vec{r}, t)\, d^3r \quad ,$$

which leads to

$$[p_{\vec{k}}(t), q_{\vec{k}'}(t)] = \frac{1}{\Omega} \int e^{i\vec{k}\cdot\vec{r} - i\vec{k}'\cdot\vec{r}'} [\Pi(\vec{r},t), \phi(\vec{r}',t)]\, d^3r\, d^3r'.$$

By using (2.5) and (2.22) we find

$$[p_{\vec{k}}(t), q_{\vec{k}'}(t)] = -i\delta_{\vec{k},\vec{k}'} \quad . \qquad (2.23)$$

Likewise, by using (2.6) we have

$$[q_{\vec{k}}(t), q_{\vec{k}'}(t)] = [p_{\vec{k}}(t), p_{\vec{k}'}(t)] = 0 \quad . \qquad (2.24)$$

These expressions enable us to derive the commutators between the $a_{\vec{k}}$'s and $a^{\dagger}_{\vec{k}}$'s. From (2.16) and (2.18) it follows that

$$[a_{\vec{k}}(t),\ a^{\dagger}_{\vec{k}'}(t)] = \delta_{\vec{k},\vec{k}'} \ ,$$

$$[a_{\vec{k}}(t),\ a_{\vec{k}'}(t)] = [a^{\dagger}_{\vec{k}}(t),\ a^{\dagger}_{\vec{k}'}(t)] = 0 \ . \qquad (2.25)$$

Remarks

1. We may reverse the above proof by starting from the commutation relations (2.25) between the $a_{\vec{k}}$'s and $a^{\dagger}_{\vec{k}}$'s, and then establishing the commutation relations (2.5) and (2.6) between ϕ and Π: From (2.25) and the Fourier expansions (2.20) and (2.21), we can directly derive (2.6) and

$$[\Pi(\vec{r}, t),\ \phi(\vec{r}', t)] = -i \sum \frac{1}{\Omega} e^{i\vec{k}\cdot(\vec{r}-\vec{r}')} \ . \qquad (2.26)$$

Furthermore, we observe that for $\vec{r}$ and $\vec{r}'$ within the volume Ω, the function $\delta^3(\vec{r}-\vec{r}')$ may be expanded in terms of the Fourier series:

$$\delta^3(\vec{r}-\vec{r}') = \sum_{\vec{k}} C_{\vec{k}}(\vec{r}') \frac{1}{\sqrt{\Omega}} e^{i\vec{k}\cdot\vec{r}} \ .$$

Because of the orthonormality relation (2.22), we have

$$C_{\vec{k}}(\vec{r}') = \int \frac{d^3r}{\sqrt{\Omega}} e^{-i\vec{k}\cdot\vec{r}} \delta^3(\vec{r}-\vec{r}') = \frac{1}{\sqrt{\Omega}} e^{-i\vec{k}\cdot\vec{r}'} \ ,$$

and therefore

$$\delta^3(\vec{r}-\vec{r}') = \sum_{\vec{k}} \frac{1}{\Omega} e^{i\vec{k}\cdot(\vec{r}-\vec{r}')} \ . \qquad (2.27)$$

Substituting (2.27) into (2.26), we obtain for $\vec{r}$ and $\vec{r}'$ within the volume Ω

$$[\Pi(\vec{r}, t),\ \phi(\vec{r}', t)] = -i\delta^3(\vec{r}-\vec{r}')$$

which is (2.5). So far, $\vec{r}$ and $\vec{r}'$ are restricted to points within Ω. In the limit $\Omega \to \infty$, the above expression becomes valid everywhere.

2. When $\Omega \to \infty$, we may replace the sum over $\vec{k}$ vectors by an integration:

$$\frac{1}{\Omega} \sum_{\vec{k}} \to \frac{1}{8\pi^3} \int d^3k \ . \qquad (2.28)$$

This can be proved by using (2.13) and observing that, in accordance with (2.14), the parameter ℓ_i runs over the discrete values $\cdots, -2, -1, 0, 1, 2, \cdots$ with a spacing

$$\Delta \ell_i = 1 \ ; \tag{2.29}$$

therefore, the corresponding variation in k_i is

$$\Delta k_i = \frac{2\pi}{L_i} \Delta \ell_i = \frac{2\pi}{L_i} \ . \tag{2.30}$$

Because of (2.29), (2.30) and $\Omega = L_1 L_2 L_3$ we may write

$$\frac{1}{\Omega} \sum_{\vec{k}} = \frac{1}{\Omega} \sum_{\vec{k}} \Delta \ell_1 \Delta \ell_2 \Delta \ell_3 = \frac{1}{8\pi^3} \sum_{\vec{k}} \Delta k_1 \Delta k_2 \Delta k_3 \ .$$

When $L_i \to \infty$, we have $\Delta k_i \to 0$. Hence the above expression leads to (2.28).

In the same limit $\Omega \to \infty$, Eq. (2.27) becomes

$$\delta^3(\vec{r} - \vec{r}') = \frac{1}{8\pi^3} \int e^{i\vec{k}\cdot(\vec{r}-\vec{r}')} d^3k \ . \tag{2.31}$$

While (2.27) is valid only for $\vec{r}$ and $\vec{r}'$ within the volume Ω, the above formula is valid for arbitrary $\vec{r}$ and $\vec{r}'$.

3. We note that the validity of the Fourier expansion (2.20)–(2.21) and the commutation relations (2.25) is independent of the detailed form of the Hamiltonian. Thus in (2.4) the function $V(\phi)$ can be of arbitrary form. If $V(\phi)$ is a quadratic function of ϕ, then it is a free-field theory, otherwise not. Furthermore, the parameter m in (2.17) is as yet completely arbitrary.

This situation is analogous to the quantum mechanics of a finite system of particles, as discussed in Chapter 1. There, the choice of the generalized coordinates $q_i(t)$ and the generalized momenta $p_i(t)$ can also be made independently of the detailed form of the interaction potential between particles.

2.3 Hilbert Space (Free or Interacting Fields)

Without any loss of generality, we can write

$$V(\phi) = \tfrac{1}{2} m^2 \phi^2 + \mathcal{H}_{int}(\phi)$$

where m is the same parameter introduced in (2.17) and $\mathcal{H}_1$ is simply defined to be

$$\mathcal{H}_{int} = V(\phi) - \tfrac{1}{2} m^2 \phi^2 \quad . \tag{2.32}$$

Thus the Hamiltonian (2.4) can be written as

$$H = H_0 + H_{int} \tag{2.33}$$

where

$$H_0 = \int \mathcal{H}_0 \, d^3r \quad , \quad H_{int} = \int \mathcal{H}_{int} \, d^3r \tag{2.34}$$

and

$$\mathcal{H}_0 = \tfrac{1}{2}(\Pi^2 + (\nabla\phi)^2 + m^2\phi^2) \quad .$$

Through partial integration H_0 becomes

$$H_0 = \tfrac{1}{2} \int [\Pi^2 + \phi(-\nabla^2 + m^2)\phi] \, d^3r \quad . \tag{2.35}$$

By using (2.17) and (2.20) we find

$$(-\nabla^2 + m^2)\phi = \sum_{\vec{k}} \frac{\omega^2}{\sqrt{2\omega\Omega}} [a_{\vec{k}}(t) e^{i\vec{k}\cdot\vec{r}} + a^\dagger_{\vec{k}}(t) e^{-i\vec{k}\cdot\vec{r}}] \quad .$$

Upon substituting this expression and (2.21) into (2.35) we have

$$H_0 = \tfrac{1}{2} \sum_{\vec{k}} \omega (a_{\vec{k}} a^\dagger_{\vec{k}} + a^\dagger_{\vec{k}} a_{\vec{k}}) = \sum_{\vec{k}} \omega (a^\dagger_{\vec{k}} a_{\vec{k}} + \tfrac{1}{2}) \quad . \tag{2.36}$$

From the example discussed in Section 1.2 we know that the eigenvalues of

$$N_{\vec{k}} \equiv a^\dagger_{\vec{k}} a_{\vec{k}} \tag{2.37}$$

are $n_{\vec{k}} = 0, 1, 2, \cdots$. Hence the operator $\sum \omega a^\dagger_{\vec{k}} a_{\vec{k}}$ is one that is bounded from below, but not from above. The totality of its eigenvectors forms a complete set, and it spans the entire Hilbert space of this system. These eigenvectors, properly normalized, are

$$|0\rangle, \; a^\dagger_{\vec{k}} |0\rangle, \; a^\dagger_{\vec{k}} a^\dagger_{\vec{k}'} |0\rangle \text{ if } \vec{k} \neq \vec{k}', \; \frac{1}{\sqrt{2}} (a^\dagger_{\vec{k}})^2 |0\rangle, \cdots \tag{2.38}$$

where the state $|0>$ satisfies

$$a_{\vec{k}} |0> = 0 \qquad \text{for all } \vec{k} . \tag{2.39}$$

The lowest-energy state of H_0 is $|0>$, which will therefore be called the vacuum state of H_0. Similarly, we may call $a^\dagger_{\vec{k}} |0>$ the corresponding one-particle state and $a^\dagger_{\vec{k}} a^\dagger_{\vec{k}'} |0>$ the two-particle state, etc. Because $a^\dagger_{\vec{k}}$ commutes with $a^\dagger_{\vec{k}'}$, we have

$$a^\dagger_{\vec{k}} a^\dagger_{\vec{k}'} |0> = a^\dagger_{\vec{k}'} a^\dagger_{\vec{k}} |0> . \tag{2.40}$$

Consequently, these particles automatically satisfy Bose statistics.

Remarks

1. As emphasized before, so far the parameter m in (2.17) can be arbitrarily chosen. Different m values give different $a_{\vec{k}}$'s and $a^\dagger_{\vec{k}}$'s in the expansion (2.20)-(2.21), and therefore different basis vectors (2.38) of the same Hilbert space. Consequently, a change in m implies a canonical transformation between the $a_{\vec{k}}$'s and the $a^\dagger_{\vec{k}}$'s.

2. If in (2.33) H_{int} is 0, i.e., the field is free, the corresponding Hamiltonian becomes

$$H = H_0 = \tfrac{1}{2} \int [\Pi^2 + (\nabla\phi)^2 + m^2 \phi^2] \, d^3r . \tag{2.41}$$

Now, set the parameter m in (2.17) to be the same one as above. We have in accordance with (2.36)

$$H = H_0 = \sum_{\vec{k}} \omega (a^\dagger_{\vec{k}} a_{\vec{k}} + \tfrac{1}{2}) . \tag{2.42}$$

By using Heisenberg's equation (1.9) and setting $O(t) = a_{\vec{k}}$ and $a^\dagger_{\vec{k}}$, we find

$$a_{\vec{k}}(t) \propto e^{-i\omega t} \quad \text{and} \quad a^\dagger_{\vec{k}}(t) \propto e^{i\omega t} , \tag{2.43}$$

which is valid for a free-field system. Since a change of the Hamiltonian from $H \to H +$ a constant does not alter the dynamics of the system, we may replace (2.42) by

$$H = H_0 = \sum_{\vec{k}} \omega\, a^{\dagger}_{\vec{k}}\, a_{\vec{k}} \quad . \tag{2.44}$$

Consequently, the energy of the vacuum state defined by (2.39) becomes zero.

3. If $H_{int} \neq 0$, then the time variation of $a_{\vec{k}}(t)$ and $a^{\dagger}_{\vec{k}}(t)$ will in general be much more complicated than (2.43). Suppose that there is no bound state in the system. Because of Lorentz invariance, the spectrum of the total Hamiltonian $H = H_0 + H_{int}$ must be given by (apart from an additive term which can be chosen to be zero)

$$\sum n_k\, \omega_{phys} \tag{2.45}$$

where $\omega_{phys} = (\vec{k}^2 + m^2_{phys})^{\frac{1}{2}}$ and $n_k = 0, 1, 2, \cdots$. By definition, m_{phys} is the observed mass of the physical particle in the system. It is convenient to choose in H_0 the parameter m to be m_{phys}; in this case the spectrum of H_0, given by (2.44), is identical to that of H, given by (2.45). As will be discussed in Chapter 5, this choice

$$m = m_{phys} \tag{2.46}$$

in H_0 brings great convenience to making the perturbation series expansion in powers of H_{int}.

4. The discussion above can readily be extended to a system of n Hermitian fields $\phi_1, \phi_2, \cdots, \phi_n$. The corresponding Lagrangian density can be written as

$$\mathcal{L} = -\sum_{i=1}^{n} \left[\tfrac{1}{2} \left(\frac{\partial \phi_i}{\partial x_\mu}\right)^2 + \tfrac{1}{2} m_i^2\, \phi_i^2 \right] - V(\phi_i) \tag{2.47}$$

where $\phi_i = \phi_i^{\dagger}$ and $i = 1, 2, \cdots, n$. In the case that $n = 2$ and $m_1 = m_2 = m$, we may express the above Lagrangian in terms of the complex field

$$\phi = \frac{1}{\sqrt{2}} (\phi_1 + i\, \phi_2) \tag{2.48}$$

and its Hermitian conjugate

$$\phi^\dagger = \frac{1}{\sqrt{2}} (\phi_1 - i\phi_2) \quad . \tag{2.49}$$

Accordingly, (2.47) becomes

$$\mathcal{L} = -\frac{\partial \phi^\dagger}{\partial x_\mu} \frac{\partial \phi}{\partial x_\mu} - m^2 \phi^\dagger \phi - V(\phi^\dagger, \phi) \quad . \tag{2.50}$$

Problem 2.1 Show that for a free Hermitian field ϕ

$$[\phi(\vec{r}, t), \phi(\vec{r}', t')] = -i D(x-x')$$

where $x = (\vec{r}, it)$, $x' = (\vec{r}', it')$,

$$D(x-x') = (2\pi)^{-3} \int d^3k \, e^{i\vec{k}\cdot(\vec{r}-\vec{r}')} \omega^{-1} \sin \omega(t-t')$$

and

$$\omega = \sqrt{\vec{k}^2 + m^2} \quad .$$

Furthermore, prove that $D(x)$ satisfies

(i) $\left(-\frac{\partial^2}{\partial t^2} + \nabla^2 - m^2\right) D(x) = 0$,

(ii) $D(x) = 0$ at $t = 0$

and (iii) $\dot{D}(x) = \delta^3(\vec{r})$ at $t = 0$.

Reference

G. Wentzel, Quantum Theory of Fields (New York, Interscience Publishers, Inc., 1949).

Chapter 3

THE SPIN-1/2 FIELD

3.1 Mathematical Preliminaries

We first introduce three 2 X 2 Pauli matrices

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which satisfy (3.1)

$$\tau_i = \tau_i^{\dagger} \ ,$$

$$[\tau_i, \tau_j] = \tau_i \tau_j - \tau_j \tau_i = 2i \epsilon_{ijk} \tau_k \ , \tag{3.2}$$

$$\{\tau_i, \tau_j\} \equiv \tau_i \tau_j + \tau_j \tau_i = 2\delta_{ij} \tag{3.3}$$

where δ_{ij} is the Kronecker symbol used before, and

$$\epsilon_{ijk} = \begin{cases} 1, & \text{if } ijk \text{ is an even permutation of } 1, 2, 3, \\ -1, & \text{if } ijk \text{ is an odd permutation} \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$

Quite often, we use the vector notation

$$\vec{\tau} = (\tau_1, \tau_2, \tau_3) \ .$$

Throughout this book, we denote the commutator and anticommutator between two matrices a and b by $[a, b] = ab - ba$ and $\{a, b\} = ab + ba$ respectively.

Next we introduce the definition of the direct product $A \times B$ of an $n \times n$ matrix $A = (A_{aa'})$ times an $m \times m$ matrix $B = (B_{bb'})$:

$$(A \times B)_{ab,\, a'b'} \equiv A_{aa'}\, B_{bb'} \tag{3.5}$$

where the subscripts a, a' can vary from 1 to n and the subscripts b, b' vary from 1 to m. Thus the matrix $A \times B$ is of dimension $nm \times nm$. One can verify readily that if matrices A and C are of the same dimension $n \times n$ and if matrices B and D are also of the same dimension $m \times m$, then

$$(A \times B) \cdot (C \times D) = (A \cdot C) \times (B \cdot D) \ ,$$

where the dot denotes the usual matrix multiplication.

The Dirac matrices $\vec{\sigma}$ and $\vec{\rho}$ are 4×4 matrices which can be expressed as the direct product between the 2×2 Pauli matrices (3.1) and the 2×2 unit matrix I:

$$\vec{\sigma} \equiv \vec{\tau} \times I \ , \quad \vec{\rho} \equiv I \times \vec{\tau} \ . \tag{3.6}$$

Consequently, we have

$$\vec{\sigma} = \begin{pmatrix} \vec{\tau} & 0 \\ 0 & \vec{\tau} \end{pmatrix} \ , \tag{3.7}$$

and $\vec{\rho} = (\rho_1 , \rho_2 , \rho_3)$ is given by

$$\rho_1 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \quad \rho_2 = \begin{pmatrix} 0 & -iI \\ iI & 0 \end{pmatrix}, \quad \rho_3 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} . \tag{3.8}$$

The $\vec{\sigma}$ and $\vec{\rho}$ matrices satisfy the following relations:

$$\begin{aligned} &\rho_i = \rho_i^{\dagger} \ , \qquad \sigma_i = \sigma_i^{\dagger} \ , \\ &[\rho_i , \rho_j] = 2i\, \epsilon_{ijk}\, \rho_k \ , \quad [\sigma_i , \sigma_j] = 2i\, \epsilon_{ijk}\, \sigma_k \ , \\ &\{\rho_i , \rho_j\} = \{\sigma_i , \sigma_j\} = 2\delta_{ij} \ , \\ &[\rho_i , \sigma_j] = 0 \ . \end{aligned} \tag{3.9}$$

Furthermore, we shall define

$$\vec{\alpha} \equiv \rho_1 \vec{\sigma} \ , \quad \beta \equiv \rho_3 \ , \tag{3.10}$$

$$\gamma_i = -i\beta\, \alpha_i = \rho_2\, \sigma_i \quad \text{and} \quad \gamma_4 = \beta = \rho_3 \ . \tag{3.11}$$

These matrices satisfy

$$\begin{aligned} &a_i = a_i^\dagger \ , \quad \beta = \beta^\dagger \ , \quad \gamma_\mu = \gamma_\mu^\dagger \ , \\ &[a_i , a_j] = [\gamma_i , \gamma_j] = 2i\,\epsilon_{ijk}\,\sigma_k \ , \\ &\{a_i , a_j\} = 2\delta_{ij} \ , \quad \{a_i , \beta\} = 0 \end{aligned} \tag{3.12}$$

and

$$\{\gamma_\mu , \gamma_\nu\} = 2\delta_{\mu\nu} \ .$$

In the above, as well as later on, all Roman subscripts i, j, k vary from 1 to 3 and all Greek subscripts μ, ν, λ vary from 1 to 4.

3.2 Free Field

The Lagrangian density of a free spin-$\frac{1}{2}$ field is

$$\mathcal{L}_{\text{free}} = -\psi^\dagger \gamma_4 \left(\gamma_\mu \frac{\partial}{\partial x_\mu} + m\right)\psi \tag{3.13}$$

where ψ is a 4×1 column matrix. (In the quantum theory, each of its matrix elements is a Hilbert space operator.)

If ψ were a classical field, then from the variational principle

$$\delta \int \mathcal{L}\, d^4x = 0 \ ,$$

we have

$$\left(\gamma_\mu \frac{\partial}{\partial x_\mu} + m\right)\psi = 0 \tag{3.14}$$

which, because of (3.11), can also be written as

$$\left(-i\vec{\alpha}\cdot\vec{\nabla} + \beta m\right)\psi = i\dot{\psi} \ . \tag{3.15}$$

We observe that Eq. (3.13) can be written as

$$\mathcal{L} = i\psi^\dagger\dot{\psi} + \cdots \tag{3.16}$$

where the $\cdots$ term does not contain $\dot{\psi}$. Let ψ_λ be the λ^{th} component of the 4×1 matrix ψ, where $\lambda = 1, 2, 3, 4$. By regarding $\psi_\lambda(\vec{r}, t)$ as a generalized coordinate, we see that its conjugate momentum is, on account of (3.16)

$$\mathcal{P}_\lambda(\vec{r}, t) = \frac{\partial \mathcal{L}}{\partial \dot{\psi}_\lambda} = i\psi_\lambda^\dagger(\vec{r}, t) \ . \tag{3.17}$$

The corresponding Hamiltonian density is

$$\mathcal{H}_{free} = \mathcal{P}_\lambda \dot{\psi}_\lambda - \mathcal{L}_{free} = \psi^\dagger \gamma_4 (\gamma_i \frac{\partial}{\partial x_i} + m)\psi \quad ,$$

which, because of (3.10)-(3.11), can also be written as

$$\mathcal{H}_{free} = \psi^\dagger (\frac{\vec{\alpha}\cdot\vec{\nabla}}{i} + \beta m)\psi \quad . \tag{3.18}$$

3.3 Quantization (Free or Interacting Fields)

We first generalize the above discussion to systems with interactions. The Lagrangian density is now given by

$$\mathcal{L} = \mathcal{L}_{free} + \mathcal{L}_{int} \tag{3.19}$$

where $\mathcal{L}_{free}$ remains given by (3.13). If the system consists of only the ψ field, then $\mathcal{L}_{int} = \mathcal{L}_{int}(\psi, \psi^\dagger)$; if there are additional fields, such as ϕ, then $\mathcal{L}_{int} = \mathcal{L}_{int}(\psi, \psi^\dagger, \phi, \dot{\phi})$. We shall assume that $\mathcal{L}_{int}$ does not contain $\partial\psi/\partial x_\mu$. Consequently the conjugate momentum of $\psi(\vec{r}, t)$ remains given by (3.17). The corresponding Hamiltonian density can now be written as

$$\mathcal{H} = \mathcal{H}_{free} + \mathcal{H}_{int} \tag{3.20}$$

where $\mathcal{H}_{free}$ is given by (3.18) and $\mathcal{H}_{int} = -\mathcal{L}_{int}(\psi, \psi^\dagger)$ if the system consists only of the ψ field. If there are additional fields such as ϕ, then $\mathcal{H}_{int} = \mathcal{H}_{int}(\psi, \psi^\dagger, \phi, \Pi)$, where Π is the conjugate momentum of ϕ.

Now we turn to the quantization problem. Following Jordan and Wigner, the quantization of a spin-$\frac{1}{2}$ field differs from that of an integer-spin field by the replacement of all equal-time commutation relations by anticommutation relations; i.e.,

$$\{\psi_\mu(\vec{r}, t), \mathcal{P}_\lambda(\vec{r}', t)\} = i\delta^3(\vec{r} - \vec{r}')\delta_{\mu\lambda} \quad ,$$

$$\{\psi_\mu(\vec{r}, t), \psi_\lambda(\vec{r}', t)\} = \{\mathcal{P}_\mu(\vec{r}, t), \mathcal{P}_\lambda(\vec{r}', t)\} = 0 \quad .$$

Because of (3.17) these relations can also be written as

$$\{\psi_\mu(\vec{r}, t),\ \psi_\lambda^\dagger(\vec{r}', t)\} = \delta^3(\vec{r} - \vec{r}')\,\delta_{\mu\lambda}\ , \tag{3.21}$$

$$\{\psi_\mu(\vec{r}, t),\ \psi_\lambda(\vec{r}', t)\} = \{\psi_\mu^\dagger(\vec{r}, t),\ \psi_\lambda^\dagger(\vec{r}', t)\} = 0\ . \tag{3.22}$$

The equation of motion remains given by Heisenberg's equation (1.9):

$$[\int \mathcal{H} d^3r,\ O(t)] = -i\dot{O}(t)\ .$$

In the case of a free field, by setting $H = \int \mathcal{H}_{\text{free}}\, d^3r$ and $O(t) = \psi(\vec{r}, t)$ we have

$$[\int \mathcal{H}_{\text{free}}\, d^3r,\ \psi] = -i\dot{\psi}\ . \tag{3.23}$$

On account of (3.21) - (3.22), we have, for any 4×4 matrix Γ whose matrix elements are c. numbers,

$$\begin{aligned}
&[\int \psi^\dagger(\vec{r}', t)\,\Gamma\,\psi(\vec{r}', t)\, d^3r',\ \psi_\lambda(\vec{r}, t)] \\
&= \int d^3r' (\psi_\mu^\dagger(\vec{r}', t)\ \Gamma_{\mu\nu}\,\psi_\nu(\vec{r}', t)\,\psi_\lambda(\vec{r}, t) \\
&\qquad - \psi_\lambda(\vec{r}, t)\,\psi_\mu^\dagger(\vec{r}', t)\ \Gamma_{\mu\nu}\,\psi_\nu(\vec{r}', t)) \\
&= \int d^3r' (-\psi_\mu^\dagger(\vec{r}', t)\ \Gamma_{\mu\nu}\,\psi_\lambda(\vec{r}, t)\,\psi_\nu(\vec{r}', t) \\
&\qquad - \psi_\lambda(\vec{r}, t)\,\psi_\mu^\dagger(\vec{r}', t)\ \Gamma_{\mu\nu}\,\psi_\nu(\vec{r}', t)) \\
&= -\int d^3r'\ \delta_{\mu\lambda}\,\Gamma_{\mu\nu}\,\delta^3(\vec{r} - \vec{r}')\,\psi_\nu(\vec{r}', t) = (-\Gamma\psi(\vec{r}, t))_\lambda\ .
\end{aligned} \tag{3.24}$$

Consequently, for the free field, (3.23) gives the same equation of motion (3.15) for the operator ψ :

$$(-i\vec{\alpha}\cdot\vec{\nabla} + \beta m)\psi = i\dot{\psi}\ .$$

Likewise, we can show that in the case of interacting fields Heisenberg's equation also leads to the same equations of motion as those given by the variational principle $\delta \int \mathcal{L}\, d^4x = 0$.

Exercise. Show that

$$\begin{aligned}
&[\int \psi^\dagger(\vec{r}', t)\ \Gamma\psi(\vec{r}', t)\, d^3r',\ \int \psi^\dagger(\vec{r}, t)\,\Gamma'\psi(\vec{r}, t)\, d^3r] \\
&\quad = \int \psi^\dagger(\vec{r}, t)\,[\Gamma, \Gamma']\,\psi(\vec{r}, t)\, d^3r
\end{aligned} \tag{3.24a}$$

where Γ and Γ' are both 4×4 matrices whose matrix elements are c. numbers.

3.4 Fourier Expansion (Free or Interacting Fields)

Just as in (2.11)-(2.12), at any given time t, the operator $\psi(\vec{r}, t)$ can be expanded in terms of the Fourier series:

$$\psi(\vec{r}, t) = \sum_{\vec{p}} S_{\vec{p}}(t) \frac{e^{i\vec{p}\cdot\vec{r}}}{\sqrt{\Omega}} \quad . \tag{3.25}$$

In the above expression $S_{\vec{p}}(t)$ is, like $\psi(\vec{r}, t)$, a 4×1 matrix with its matrix elements the Hilbert-space operators. The only difference is that, unlike $\psi(\vec{r}, t)$, $S_{\vec{p}}(t)$ is independent of $\vec{r}$.

Let us regard the 4×1 column matrices as vectors in a 4-dimensional space, called spinor space. For a given $\vec{p}$ it is convenient to introduce the following set of c. number basis vectors $u_{\vec{p},s}$ and $v_{-\vec{p},s}$ in the spinor space. These vectors satisfy

$$(\vec{\alpha}\cdot\vec{p} + \beta m) \begin{cases} u_{\vec{p},s} \\ v_{-\vec{p},s} \end{cases} = E_p \begin{cases} u_{\vec{p},s} \\ -v_{-\vec{p},s} \end{cases} \tag{3.26}$$

and

$$\vec{\sigma}\cdot\hat{p} \begin{cases} u_{\vec{p},s} \\ v_{-\vec{p},s} \end{cases} = 2s \begin{cases} u_{\vec{p},s} \\ v_{-\vec{p},s} \end{cases} \tag{3.27}$$

where

$$E_p = \sqrt{\vec{p}^2 + m^2} > 0 \quad , \tag{3.28}$$

$$\hat{p} = \frac{\vec{p}}{|\vec{p}|}$$

and the parameter

$$s = \pm \tfrac{1}{2} \tag{3.29}$$

is called helicity, whose physical significance will be discussed in Section 3.7. We shall normalize these vectors so that

$$u^{\dagger}_{\vec{p},s}\, u_{\vec{p},s} = v^{\dagger}_{-\vec{p},s}\, v_{-\vec{p},s} = 1 \quad . \tag{3.30}$$

At o given $\vec{p}$, the 4×4 motrices $(\vec{\alpha} \cdot \vec{p} + \beta m)$ ond $\vec{\sigma} \cdot \hat{p}$ are both Hermition. According to (3.26) and (3.27) the four vectors $u_{\vec{p},s}$ and $v_{-\vec{p},s}$ with $s = \pm \frac{1}{2}$ are eigenvectors of these two Hermition motrices with different eigenvalues. Consequently, these four vectors ore orthogonol to eoch other; becouse of (3.30) they form o complete orthonormal set of basis vectors in the spinor space. The $S_{\vec{p}}(t)$ in Eq. (3.25) can be exponded in terms of this set of basis vectors:

$$S_{\vec{p}}(t) = \sum_{s=\pm\frac{1}{2}} (o_{\vec{p},s}(t)\, u_{\vec{p},s} + b^{\dagger}_{-\vec{p},s}(t)\, v_{-\vec{p},s}) \tag{3.31}$$

where the coefficients $o_{\vec{p},s}(t)$ ond $b^{\dagger}_{-\vec{p},s}(t)$ are Hilbert-space operators. Combining (3.25) ond (3.31) we have

$$\psi(\vec{r}, t) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{p},s} (a_{\vec{p},s}(t)\, u_{\vec{p},s}\, e^{i\vec{p}\cdot\vec{r}} + b^{\dagger}_{\vec{p},s}(t)\, v_{\vec{p},s}\, e^{-i\vec{p}\cdot\vec{r}}) \,. \tag{3.32}$$

Its Hermition conjugate is

$$\psi^{\dagger}(\vec{r}, t) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{p},s} (o^{\dagger}_{\vec{p},s}(t)\, u^{\dagger}_{\vec{p},s}\, e^{-i\vec{p}\cdot\vec{r}} + b_{\vec{p},s}(t)\, v^{\dagger}_{\vec{p},s}\, e^{i\vec{p}\cdot\vec{r}}) \,. \tag{3.33}$$

From the onticommutotion relotions (3.21)–(3.22) we con reodily verify

$$\begin{aligned}
&\{a_{\vec{p},s}(t),\ a^{\dagger}_{\vec{p}',s'}(t)\} = \{b_{\vec{p},s}(t),\ b^{\dagger}_{\vec{p}',s'}(t)\} = \delta_{\vec{p},\vec{p}'}\, \delta_{s,s'} \,, \\
&\{o_{\vec{p},s}(t),\ a_{\vec{p}',s'}(t)\} = \{b_{\vec{p},s}(t),\ b_{\vec{p}',s'}(t)\} = 0 \,, \\
&\{o_{\vec{p},s}(t),\ b_{\vec{p}',s'}(t)\} = \{o_{\vec{p},s}(t),\ b^{\dagger}_{\vec{p}',s'}(t)\} = 0 \,.
\end{aligned} \tag{3.34}$$

As noted before, ony function of $\vec{r}$ con be expanded in terms of the complete set $\{\frac{1}{\sqrt{\Omega}} e^{i\vec{p}\cdot\vec{r}}\}$, ond ony 4×1 column motrix in the spinor space con be expanded in terms of the four orthonormol bosis vectors: $u_{\vec{p},\pm\frac{1}{2}}$ ond $v_{-\vec{p},\pm\frac{1}{2}}$ where $\vec{p}$ is fixed. Consequently the expansion (3.32) is volid for the free, os well as the interocting, field.

For the free-field cose, the Homiltonion density is given by

$$\mathcal{H} = \mathcal{H}_{\text{free}} = \psi^\dagger (-i\vec{\alpha}\cdot\vec{\nabla} + \beta m)\,\psi \quad . \tag{3.35}$$

By substituting (3.32) into the above expression, we obtain

$$H = H_{\text{free}} = \int \mathcal{H}_{\text{free}}\, d^3r = \sum_{\vec{p},s} (a^\dagger_{\vec{p},s}\, a_{\vec{p},s} - b_{\vec{p},s}\, b^\dagger_{\vec{p},s})\, E_p$$

$$= \sum_{\vec{p},s} (a^\dagger_{\vec{p},s}\, a_{\vec{p},s} + b^\dagger_{\vec{p},s}\, b_{\vec{p},s} - 1)\, E_p \,. \tag{3.36}$$

Since a change $H \to H +$ a constant does not alter the dynamics of the system, we may drop the -1 inside the parentheses in the above formula. The Hamiltonian (3.36) is then replaced by

$$H = H_{\text{free}} = \sum_{\vec{p},s} (a^\dagger_{\vec{p},s}\, a_{\vec{p},s} + b^\dagger_{\vec{p},s}\, b_{\vec{p},s})\, E_p \quad . \tag{3.37}$$

By using Heisenberg's equation (1.9) and by setting $O(t) = a_{\vec{p},s}(t)$, we have for the free-field case

$$-i\,\dot{a}_{\vec{p},s}(t) = [H,\, a_{\vec{p},s}(t)] = -E_p\, a_{\vec{p},s}(t) \quad ,$$

and therefore

$$a_{\vec{p},s}(t) \propto e^{-iE_p t} \quad . \tag{3.38}$$

Likewise we can derive

$$b_{\vec{p},s}(t) \propto e^{-iE_p t} \quad . \tag{3.39}$$

In the case of an interacting field, the time-dependence of $a_{\vec{p},s}(t)$ and $b_{\vec{p},s}(t)$ will in general be more complicated.

Exercise. Show that

$$\sum_s u_{\vec{p},s}\, u^\dagger_{\vec{p},s}\, \beta = \frac{\not{p} + m}{2p_0} \quad ,$$

$$\sum_s v_{\vec{p},s}\, v^\dagger_{\vec{p},s}\, \beta = \frac{\not{p} - m}{2p_0} \quad ,$$

$$\tfrac{1}{4}\,\text{trace}\,(\not{A}\not{B}) = -(A\cdot B) \equiv A_0 B_0 - \vec{A}\cdot\vec{B} \quad ,$$

$$\tfrac{1}{4}\,\text{trace}\,(\not{A}\not{B}\not{C}\not{D}) = (A\cdot B)(C\cdot D) - (A\cdot C)(B\cdot D) + (A\cdot D)(B\cdot C)$$

where A, B, C and D can be any c. number 4-vectors,

$\not{A} = -i\gamma_\mu A_\mu$, $A_4 = iA_0$, $\not{B} = -i\gamma_\mu B_\mu$, $B_4 = iB_0$, etc.

3.5 Hilbert Space (Free or Interacting Fields)

Just as in (2.38), the Hilbert space is spanned by the set of orthonormal basis vectors:

$$|0>\ ,\quad a^\dagger_{\vec{p},s}|0>\ ,\quad b^\dagger_{\vec{p},s}|0>\ ,\quad a^\dagger_{\vec{p},s}a^\dagger_{\vec{p}',s'}|0>\ ,$$

$$a^\dagger_{\vec{p},s}b^\dagger_{\vec{p}',s'}|0>\ ,\quad b^\dagger_{\vec{p},s}b^\dagger_{\vec{p}',s'}|0>\ ,\ \cdots \tag{3.40}$$

where the state vector $|0>$ satisfies

$$a_{\vec{p},s}|0> = 0 \quad \text{and} \quad b_{\vec{p},s}|0> = 0 \tag{3.41}$$

for all $\vec{p}$ and s. In order to analyze the structure of this Hilbert space, we must first discuss some elementary algebraic properties of these anticommuting operators $a_{\vec{p},s}$, $b_{\vec{p},s}$ and their Hermitian conjugates.

(i) We first discuss the case of a single mode. Let a and $a^\dagger$ satisfy the following anticommutation relations

$$\{a, a^\dagger\} = 1\ , \tag{3.42}$$

and

$$\{a, a\} = \{a^\dagger, a^\dagger\} = 0\ ;$$

the latter can also be written as

$$a^2 = (a^\dagger)^2 = 0\ . \tag{3.43}$$

Let us define

$$N \equiv a^\dagger a\ . \tag{3.44}$$

Because of (3.42)-(3.43), we find

$$N^2 = a^\dagger a\, a^\dagger a = a^\dagger(1 - a^\dagger a)\, a = N\ ,$$

which implies that the eigenvalues of N can only be 0 or 1. Assuming that N does have an eigenstate, denoted by $|0>$, with the

eigenvalue 0 :

$$N\,|\,0\rangle = 0 \quad . \tag{3.45}$$

Then it follows that

$$Na^\dagger\,|\,0\rangle = a^\dagger a\,a^\dagger\,|\,0\rangle = a^\dagger(1-a^\dagger a)\,|\,0\rangle = a^\dagger\,|\,0\rangle \quad . \tag{3.46}$$

By designating

$$|\,1\rangle \equiv a^\dagger\,|\,0\rangle \quad , \tag{3.47}$$

we can write (3.46) as

$$N\,|\,1\rangle = |\,1\rangle \quad . \tag{3.48}$$

Thus, the existence of the eigenstate $|\,0\rangle$ implies that of $|\,1\rangle$. The converse is also true, since from (3.47) we can establish

$$|\,0\rangle = a\,|\,1\rangle \quad . \tag{3.49}$$

Therefore, both eigenstates exist. Furthermore, because of (3.43) we have

$$a^\dagger\,|\,1\rangle = 0 \quad \text{and} \quad a\,|\,0\rangle = 0 \quad . \tag{3.50}$$

These two eigenvectors $|\,0\rangle$ and $|\,1\rangle$ span a two-dimensional Hilbert space. We may represent

$$|\,0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |\,1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad . \tag{3.51}$$

The operators a , $a^\dagger$ and N can be expressed in matrix form:

$$\begin{aligned} a &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \equiv \tau_+ = \tfrac{1}{2}(\tau_1 + i\,\tau_2) \quad , \\ a^\dagger &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \equiv \tau_- = \tfrac{1}{2}(\tau_1 - i\,\tau_2) \quad , \\ N &= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \tfrac{1}{2}(1 - \tau_3) \end{aligned} \tag{3.52}$$

where τ_1 , τ_2 and τ_3 are 2×2 Pauli matrices given by (3.1). As in the case of the boson field, we call N the occupation-number operator, a the annihilation operator and $a^\dagger$ the creation operator.

(ii) Next we consider the case of two modes. There are now two

annihilation operators a_1 and a_2 ; their Hermitian conjugates form two creation operators. These operators satisfy

$$\{a_i, a_j^\dagger\} = \delta_{ij}$$
$$\text{and} \qquad \{a_i, a_j\} = \{a_i^\dagger, a_j^\dagger\} = 0 \tag{3.53}$$

where i and j can be 1 or 2. We may define

$$N_1 = a_1^\dagger a_1 \quad \text{and} \quad N_2 = a_2^\dagger a_2 \ . \tag{3.54}$$

Because of (3.53), N_1 commutes with N_2. By following an argument similar to that in case (i) we can show that the eigenvalue of each N_i can be 0 or 1. Thus, the eigenvalues of the set (N_1, N_2) can be (0, 0), (1, 0), (0, 1) and (1, 1). By regarding the corresponding eigenstates as the basis vectors, we form a four-dimensional Hilbert space. In this space the matrix representations of N_1 and N_2 are

$$N_1 = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 0 & \\ & & & 1 \end{pmatrix} \ , \quad N_2 = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \ . \tag{3.55}$$

The corresponding matrices for a_i and $a_i^\dagger$ may be given by the following direct products:

$$\begin{aligned} a_1 &= \tau_+ \times I \ , \quad a_1^\dagger = \tau_- \times I \ , \\ a_2 &= \tau_3 \times \tau_+ \ , \quad a_2^\dagger = \tau_3 \times \tau_- \end{aligned} \tag{3.56}$$

where I is a 2×2 unit matrix. In explicit form, (3.56) can also be written as

$$\begin{aligned} a_1 &= \begin{pmatrix} \tau_+ & 0 \\ 0 & \tau_+ \end{pmatrix} \ , \quad a_1^\dagger = \begin{pmatrix} \tau_- & 0 \\ 0 & \tau_- \end{pmatrix} \ , \\ a_2 &= \begin{pmatrix} 0 & \tau_3 \\ 0 & 0 \end{pmatrix} \ , \quad a_2^\dagger = \begin{pmatrix} 0 & 0 \\ \tau_3 & 0 \end{pmatrix} \ . \end{aligned} \tag{3.57}$$

By using (3.56) one may verify directly that the anticommutation relations given by (3.53) hold. Furthermore, the matrices N_1 and

N_2 are given by (3.55).

(iii) The above considerations can be readily generalized to the case of n modes. The anticommutation relations have the same form as (3.53); i.e.,

$$\{a_i, a_j^\dagger\} = \delta_{ij} ,$$
$$\{a_i, a_j\} = \{a_i^\dagger, a_j^\dagger\} = 0 \tag{3.58}$$

except that i and j can now vary from 1 to n. By induction we can generalize (3.56) to

$$a_1 = \tau_+ \times I \times I \times \cdots \times I , \quad a_1^\dagger = \tau_- \times I \times I \times \cdots \times I ,$$
$$a_2 = \tau_3 \times \tau_+ \times I \times \cdots \times I , \quad a_2^\dagger = \tau_3 \times \tau_- \times I \times \cdots \times I ,$$
$$\cdots\cdots \tag{3.59}$$
$$a_n = \tau_3 \times \tau_3 \times \cdots \times \tau_+ , \quad a_n^\dagger = \tau_3 \times \tau_3 \times \cdots \times \tau_-$$

in which the expression for a_1 contains n - 1 factors of I, that for a_2 n - 2 factors of I, etc. These matrices in (3.59) can be easily seen to satisfy the anticommutation relations (3.58). Furthermore, the operator

$$N_i \equiv a_i^\dagger a_i$$

has the eigenvalue 0 or 1, where the subscript i can be $1, 2, \cdots, n$.

Remarks. In the general case, the Hilbert space is spanned by the basis vectors (3.40), in which $| 0 >$ is called the vacuum state of H_{free}, which is determined by (3.41). Likewise, $a^\dagger_{\vec{p},s} | 0 >$ is called a one-particle state, $b^\dagger_{\vec{p},s} | 0 >$ a one-antiparticle state, $a^\dagger_{\vec{p},s} a^\dagger_{\vec{p}',s'} | 0 >$ a two-particle state, etc. As we shall discuss in the next section, the subscripts $\vec{p}$ and s denote respectively the momentum and helicity of the particle (or antiparticle), where "helicity" means the component of angular momentum along the direction of $\vec{p}$.

Because of the anticommutation relations (3.34), we have

$$a^{\dagger}_{\vec{p},s}\, a^{\dagger}_{\vec{p}',s'} \mid 0> \;=\; - a^{\dagger}_{\vec{p}',s'}\, a^{\dagger}_{\vec{p},s} \mid 0> \tag{3.60}$$

which implies that these particles obey Fermi statistics. Thus, for example, when $\vec{p} = \vec{p}'$ and $s = s'$ the vector (3.60) is a null vector, showing that these particles do satisfy Pauli's exclusion principle.

Exercise. In the case of bosons with n modes, the commutation relations between the annihilation and creation operators are given by

$$[a_i\,,\; a_j^{\dagger}] \;=\; \delta_{ij}$$

and

$$[a_i\,,\; a_j] \;=\; [a_i^{\dagger},\; a_j^{\dagger}] \;=\; 0 \quad ,$$

where i and j vary from 1 to n. Show that the matrix forms of these operators can be written as the following direct products

$$a_1 = a \times I \times I \times \cdots \times I \,, \quad a_1^{\dagger} = a^{\dagger} \times I \times I \times \cdots \times I \,,$$

$$a_2 = I \times a \times I \times \cdots \times I \,, \quad a_2^{\dagger} = I \times a^{\dagger} \times I \times \cdots \times I \,,$$

$$\cdots\cdots$$

$$a_n = I \times I \times \cdots \times I \times a \,, \quad a_n^{\dagger} = I \times I \times \cdots \times I \times a^{\dagger}$$

where a and $a^{\dagger}$ are given by (1.31) and I is an $\infty \times \infty$ unit matrix.

3.6 Momentum and Angular Momentum Operators

Let us define the momentum operator $\vec{P}$ and the angular momentum operator $\vec{J}$ of a spin-$\frac{1}{2}$ field to be

$$\vec{P}(t) \;\equiv\; -i \int \psi^{\dagger}(\vec{r},\, t)\, \vec{\nabla}\, \psi(\vec{r},\, t)\, d^3r \tag{3.61}$$

and

$$\vec{J}(t) \;\equiv\; \int \psi^{\dagger}(\vec{r},\, t)\, (\vec{\ell} + \tfrac{1}{2}\vec{\sigma})\, \psi(\vec{r},\, t)\, d^3r \tag{3.62}$$

where

$$\vec{\ell} \;=\; -i\vec{r} \times \vec{\nabla} \quad . \tag{3.63}$$

By using (3.24), we find

$$[\, \vec{P}(t)\,,\; \psi(\vec{r},\, t)\,] \;=\; i\, \vec{\nabla}\, \psi(\vec{r},\, t) \tag{3.64}$$

and

$$[\vec{J}(t),\ \psi(\vec{r}, t)] = -(-i\vec{r}\times\vec{\nabla} + \tfrac{1}{2}\vec{\sigma})\,\psi(\vec{r}, t)\ . \tag{3.65}$$

Thus, operating on $\psi(\vec{r}, t)$, $\vec{P}(t)$ acts as an infinitesimal displacement operator and $\vec{J}(t)$ as an infinitesimal rotation operator. [See Problem 3 at the end of Chapter 10.]

By using the exercise given in Section 3.3 we can readily verify that the components of $\vec{P}(t)$ and $\vec{J}(t)$ satisfy

$$[P_i(t),\ P_j(t)] = 0\ , \tag{3.66}$$

$$[J_i(t),\ J_j(t)] = i\epsilon_{ijk}J_k(t)\ , \tag{3.67}$$

$$[P_1(t),\ J_1(t)] = [P_2(t),\ J_2(t)] = [P_3(t),\ J_3(t)] = 0\ , \tag{3.68}$$

but

$$[P_i(t),\ J_j(t)] \neq 0 \quad \text{if} \quad i \neq j\ , \tag{3.69}$$

where i and j can be 1, 2 or 3, and ϵ_{ijk} is given by (3.4). The above relations merely reflect the fact that the differential operators ∇_i and ∇_j commute,

$$[\ell_i, \ell_j] = i\epsilon_{ijk}\ell_k, \quad [\tfrac{1}{2}\sigma_i, \tfrac{1}{2}\sigma_j] = i\epsilon_{ijk}\tfrac{1}{2}\sigma_k$$

and while ∇_i commutes with $(\vec{r}\times\vec{\nabla})_i$, it does not commute with $(\vec{r}\times\vec{\nabla})_j$ for $j \neq i$.

If the system is a free spin-$\frac{1}{2}$ field, then the Hamiltonian is

$$H_{free} = \int \mathcal{H}_{free}\, d^3r \tag{3.70}$$

where $\mathcal{H}_{free}$ is given by (3.18). It is straightforward to derive

$$[\vec{P}(t),\ H_{free}] = 0$$

and

$$[\vec{J}(t),\ H_{free}] = 0\ . \tag{3.71}$$

Both $\vec{P}$ and $\vec{J}$ are therefore constants of motion. In the case of interacting fields, the total momentum consists of $\vec{P}(t)$ of the spin-$\frac{1}{2}$ field, plus the momenta of other fields. Consequently, conservation of total momentum does not imply that $\vec{P}(t)$ is a constant of motion.

Similar considerations also apply to the total angular momentum.

We may express $\vec{P}$ in terms of the Fourier components of ψ. Let us substitute (3.32)-(3.33) into (3.61). By using (2.22) and the orthonormality relations between the spinors $u_{\vec{p},s}$ and $v_{\vec{p},s}$ we obtain

$$\vec{P} = \sum_{\vec{p},s} \vec{p}(a^{\dagger}_{\vec{p},s}\, a_{\vec{p},s} - b_{\vec{p},s}\, b^{\dagger}_{\vec{p},s})$$
$$= \sum_{\vec{p},s} \vec{p}(a^{\dagger}_{\vec{p},s}\, a_{\vec{p},s} + b^{\dagger}_{\vec{p},s}\, b_{\vec{p},s} - 1) \quad .$$

Since $\sum_{\vec{p}} \vec{p} = 0$ due to symmetry, the above expression can be written as

$$\vec{P} = \sum_{\vec{p},s} \vec{p}(a^{\dagger}_{\vec{p},s}\, a_{\vec{p},s} + b^{\dagger}_{\vec{p},s}\, b_{\vec{p},s}) \quad . \tag{3.72}$$

For definiteness, we may consider ψ to be the electron field, and denote in the Hilbert space (3.40)

$$| e^{-}_{\vec{p},s} > \equiv a^{\dagger}_{\vec{p},s}(t) \,|\, 0 >$$
and
$$| e^{+}_{\vec{p},s} > \equiv b^{\dagger}_{\vec{p},s}(t) \,|\, 0 > \quad . \tag{3.73}$$

From (3.72), it follows that these states are eigenvectors of $\vec{P}$:

$$\vec{P} \,|\, 0 > = 0 \tag{3.74}$$

and

$$\vec{P} \,|\, e^{\pm}_{\vec{p},s} > = \vec{p} \,|\, e^{\pm}_{\vec{p},s} > \quad . \tag{3.75}$$

The former can also be derived by noting that the validity of all our expressions from (3.40) on is independent of the Hamiltonian H, provided that all operators are synchronized at the same time t. Thus,

we need only consider the special case $H = H_{free}$. Because of (3.71) and the fact that $|\,0\,>$ is the ground state of H_{free} with no degeneracy, (3.74) follows. Likewise, we have

$$\vec{J}\,|\,0\,> \;=\; 0 \quad . \tag{3.76}$$

Next, we shall prove

$$\vec{J}\cdot\vec{P}\;|\;e^{\pm}_{\vec{p},s}\;> \;=\; |\,\vec{p}\,|\;s\;|\;e^{\pm}_{\vec{p},s}\;> \tag{3.77}$$

which, together with (3.75), means that $\vec{p}$ is the momentum of the state and the helicity $s = \pm\frac{1}{2}$ is its component of angular momentum along $\vec{p}$. To see this, we take the Hermitian conjugates of (3.64) and (3.65) and apply them onto $|\,0\,>$. That gives, on account of (3.74) and (3.76)

$$P_k(t)\;\psi^{\dagger}(\vec{r},\,t)\;|\,0\,> \;=\; -i\,\nabla_k\,\psi^{\dagger}(\vec{r},\,t)\;|\,0\,> \quad ,$$

$$J_k(t)\;\psi^{\dagger}(\vec{r},\,t)\;|\,0\,> \;=\; -(i\,\vec{r}\times\vec{\nabla}\,\psi^{\dagger} + \tfrac{1}{2}\,\psi^{\dagger}\vec{\sigma})_k\;|\,0\,>$$

and therefore

$$\vec{J}\cdot\vec{P}\;\psi^{\dagger}(\vec{r},\,t)\;|\,0\,> \;=\; i\,\tfrac{1}{2}\,\nabla_k\,\psi^{\dagger}(\vec{r},\,t)\;\sigma_k\;|\,0\,> \,. \tag{3.78}$$

Because

$$a^{\dagger}_{\vec{p},s}(t) \;=\; \int \frac{1}{\sqrt{\Omega}}\;e^{i\vec{p}\cdot\vec{r}}\;\psi^{\dagger}(\vec{r},\,t)\;u_{\vec{p},s}\,d^3r \quad , \tag{3.79}$$

we obtain

$$\vec{J}\cdot\vec{P}\;a^{\dagger}_{\vec{p},s}\;|\,0\,> \;=\; |\,\vec{p}\,|\;s\;a^{\dagger}_{\vec{p},s}\;|\,0\,> \quad . \tag{3.80}$$

Identical considerations can be applied to $b^{\dagger}_{\vec{p},s}\;|\,0\,>$. Equation (3.77) now follows.

From (3.68) and (3.69) it follows that the component of $\vec{J}$ along $\vec{P}$ commutes with $\vec{P}$. Therefore these two operators can be diagonalized simultaneously; their eigenvalues for the states (3.73) are respectively the helicity s and the momentum $\vec{p}$. Under a space rotation, helicity is invariant because it is the scalar product of two vectors. As we shall see in Section 3.8, if the particle is of zero mass, then helicity is also invariant under a Lorentz transformation.

3.7 Phase Factor Conventions between the Spinors

In the expansion (3.31), the four orthogonal basis vectors $u_{\vec{p},s}$ and $v_{-\vec{p},s}$ in the spinor space are determined by (3.26)-(3.30) where $\vec{p}$ is a given fixed vector. Because (3.26) and (3.27) are homogeneous equations, the phase factors of these c. number spinors $u_{\vec{p},s}$ and $v_{-\vec{p},s}$ remain arbitrary. In this section we shall show that it is possible to choose their phase factors such that

$$\text{(i)} \quad v_{\vec{p},s} = \gamma_2 u^*_{\vec{p},s}, \qquad u_{\vec{p},s} = \gamma_2 v^*_{\vec{p},s}, \tag{3.81}$$

$$\text{(ii)} \quad \gamma_4 u_{\vec{p},s} = u_{-\vec{p},-s}, \qquad \gamma_4 v_{\vec{p},s} = -v_{-\vec{p},-s}, \tag{3.82}$$

and

$$\text{(iii)} \quad \sigma_2 u^*_{\vec{p},s} = e^{i\theta_{\vec{p},s}} u_{-\vec{p},s}, \qquad \sigma_2 v^*_{\vec{p},s} = e^{-i\theta_{-\vec{p},s}} v_{-\vec{p},s} \tag{3.83}$$

where

$$e^{i\theta_{\vec{p},s}} = -e^{i\theta_{-\vec{p},s}} . \tag{3.84}$$

These phase factor conventions will be adopted throughout this book.

Proof. We note that from (3.1)-(3.11) the matrices

$$\rho_1, \ \rho_3, \ \sigma_1, \ \sigma_3, \ a_1, \ a_3, \ \gamma_2$$

are real and the matrices

$$\rho_2, \ \sigma_2, \ a_2, \ \gamma_1, \ \gamma_3$$

are imaginary.

(i) Because $\gamma_2 = \rho_2 \sigma_2$, we have

$$\gamma_2 \vec{\sigma}^* \gamma_2 = -\vec{\sigma} , \qquad \gamma_2 \vec{\alpha}^* \gamma_2 = \vec{\alpha} . \tag{3.85}$$

By taking the complex conjugate of (3.26) and noting that $\vec{p}$, m, β and E_p are all real, we obtain

$$(\vec{\alpha}^* \cdot \vec{p} + \beta m)\, u^*_{\vec{p},s} = E_p\, u^*_{\vec{p},s} .$$

Through (3.85) and $\{\gamma_2, \beta\} = 0$, the above equation becomes

$$(-\vec{\alpha} \cdot \vec{p} + \beta m)(\gamma_2\, u^*_{\vec{p},s}) = -E_p\, \gamma_2\, u^*_{\vec{p},s} . \tag{3.86}$$

Likewise the complex conjugate of (3.27) is

$$(\vec{\sigma}^* \cdot \hat{p})\, u^*_{\vec{p},s} = 2s\, u^*_{\vec{p},s} ,$$

which, on account of (3.85), can also be written as

$$(-\vec{\sigma} \cdot \hat{p})\, \gamma_2\, u^*_{\vec{p},s} = 2s\, \gamma_2\, u^*_{\vec{p},s} . \tag{3.87}$$

By comparing (3.86) and (3.87) with (3.26) and (3.27) we see that $\gamma_2\, u^*_{\vec{p},s}$ and $v_{\vec{p},s}$ satisfy identical equations; therefore they must be proportional to each other. Because of the normalization condition (3.30), the proportionality constant must only be a phase factor, which can be chosen to be 1. Thus, we derive the first equation in (3.81):

$$v_{\vec{p},s} = \gamma_2\, u^*_{\vec{p},s} . \tag{3.88}$$

Since $\gamma_2 = \rho_2 \sigma_2$ is a real matrix and $\gamma_2^2 = 1$, the complex conjugate of the above equation gives

$$u_{\vec{p},s} = \gamma_2\, v^*_{\vec{p},s} .$$

(ii) Because

$$\gamma_4 \vec{\alpha}\, \gamma_4 = -\vec{\alpha} , \qquad \gamma_4 \vec{\sigma}\, \gamma_4 = \vec{\sigma} ,$$

Eq. (3.26), when multiplied by γ_4 on the left, becomes

$$(-\vec{\alpha}\cdot\vec{p}+\beta m)\,\gamma_4\,u_{\vec{p},s} = E_p\,\gamma_4\,u_{\vec{p},s} \quad . \tag{3.89}$$

Likewise, a similar multiplication onto (3.27) leads to

$$(-\vec{\sigma}\cdot\hat{p})\,\gamma_4\,u_{\vec{p},s} = -2s\,\gamma_4\,u_{\vec{p},s} \quad . \tag{3.90}$$

By comparing (3.89) and (3.90) with (3.26) and (3.27) we see that $\gamma_4\,u_{\vec{p},s}$ and $u_{-\vec{p},-s}$ satisfy identical equations; consequently they differ only in a phase factor. By choosing that phase factor to be 1, we have

$$\gamma_4\,u_{\vec{p},s} = u_{-\vec{p},-s} \quad .$$

Taking the complex conjugate and multiplying on the left by γ_2, we obtain

$$\gamma_2\,\gamma_4\,u^*_{\vec{p},s} = \gamma_2\,u^*_{-\vec{p},-s} \quad ,$$

which, because of (3.88), leads to

$$\gamma_4\,v_{\vec{p},s} = -v_{-\vec{p},-s} \quad .$$

(iii) We note that

$$\sigma_2\,\vec{\sigma}^*\,\sigma_2 = -\vec{\sigma} \quad , \quad \sigma_2\,\vec{\alpha}^*\,\sigma_2 = -\vec{\alpha} \quad .$$

By first taking the complex conjugate of (3.26) and then multiplying it on the left by σ_2, we find

$$(-\vec{\alpha}\cdot\vec{p}+\beta m)\,\sigma_2\,u^*_{\vec{p},s} = E_p\,\sigma_2\,u^*_{\vec{p},s} \quad .$$

A similar operation on (3.27) leads to

$$(-\vec{\sigma}\cdot\hat{p})\,\sigma_2\,u^*_{\vec{p},s} = 2s\,\sigma_2\,u^*_{\vec{p},s} \quad .$$

Thus $\sigma_2\,u^*_{\vec{p},s}$ and $u_{-\vec{p},s}$ satisfy identical equations, which implies

$$\sigma_2\,u^*_{\vec{p},s} \propto u_{-\vec{p},s} \quad ,$$

and therefore

$$\sigma_2\,u^*_{\vec{p},s} = e^{i\theta_{\vec{p},s}}\,u_{-\vec{p},s} \tag{3.91}$$

where the phase angle $\theta_{\vec{p},s}$ is real. Because σ_2 is an imaginary matrix, the complex conjugate of (3.91) gives

$$-\sigma_2 u_{\vec{p},s} = e^{-i\theta_{\vec{p},s}} u^*_{-\vec{p},s} , \tag{3.92}$$

which can also be written as

$$\sigma_2 u^*_{-\vec{p},s} = -e^{i\theta_{\vec{p},s}} u_{\vec{p},s} . \tag{3.93}$$

By changing the subscript $\vec{p}$ to $-\vec{p}$, we convert the above equation into

$$\sigma_2 u^*_{\vec{p},s} = -e^{i\theta_{-\vec{p},s}} u_{-\vec{p},s} . \tag{3.94}$$

A comparison between (3.91) and (3.94) gives (3.84). By using (3.81), (3.84) and (3.92), we have

$$\begin{aligned} \sigma_2 v^*_{\vec{p},s} &= \sigma_2 (\gamma_2 u^*_{\vec{p},s})^* = \gamma_2 \sigma_2 u_{\vec{p},s} \\ &= -e^{-i\theta_{\vec{p},s}} \gamma_2 u^*_{-\vec{p},s} = -e^{-i\theta_{\vec{p},s}} v_{-\vec{p},s} \\ &= e^{-i\theta_{-\vec{p},s}} v_{-\vec{p},s} . \end{aligned} \tag{3.95}$$

That completes the proof of (3.81)-(3.84).

As we shall see later, the convention (3.81) is useful for discussion of particle-antiparticle conjugation, (3.82) for parity, and (3.83) for time reversal. We note that for a given helicity s, spinors with different $\vec{p}$ in the set $\{u_{\vec{p},s}\}$ are connected through rotations and Lorentz transformations. In (3.81) we fix the relative phase factors between members of the set $\{u_{\vec{p},s}\}$ and the corresponding ones in $\{v_{\vec{p},s}\}$; likewise (3.82) relates those between $\{u_{\vec{p},s}\}$ and $\{u_{-\vec{p},-s}\}$. On the other hand, (3.83) connects the relative phase between $u_{\vec{p},s}$ and $u_{-\vec{p},s}$, which are members of the same set $\{u_{\vec{p},s}\}$. Because of continuity, the phase factor $e^{i\theta_{\vec{p},s}}$ in (3.91) cannot be 1 for all $\vec{p}$;

as shown explicitly by (3.84).

3.8 Two-component Theory

Let us first consider the case of a quantized free spin-$\frac{1}{2}$ field with $m = 0$. The equation of motion (3.14) now becomes

$$\gamma_\mu \frac{\partial}{\partial x_\mu} \psi = 0 \quad . \tag{3.96}$$

It is useful to introduce

$$\gamma_5 \equiv \gamma_1 \gamma_2 \gamma_3 \gamma_4 \quad .$$

Because of (3.12) we obtain

$$\{\gamma_5, \gamma_\mu\} = 0 \quad , \tag{3.97}$$

where $\mu = 1, 2, 3, 4$, and

$$\gamma_5^2 = 1 \quad . \tag{3.98}$$

Thus, from (3.96) we also have

$$\gamma_\mu \frac{\partial}{\partial x_\mu} (\gamma_5 \psi) = 0 \quad . \tag{3.99}$$

We may decompose

$$\psi = \psi_L + \psi_R \tag{3.100}$$

where

$$\psi_L = \tfrac{1}{2}(1 + \gamma_5)\psi \quad \text{and} \quad \psi_R = \tfrac{1}{2}(1 - \gamma_5)\psi \quad . \tag{3.101}$$

From (3.96) and (3.99) it follows that ψ_L and ψ_R separately satisfy the equation of motion; i.e.

$$\gamma_\mu \frac{\partial}{\partial x_\mu} \psi_L = 0 \quad \text{and} \quad \gamma_\mu \frac{\partial}{\partial x_\mu} \psi_R = 0 \quad . \tag{3.102}$$

In the representation (3.11), $\gamma_i = \rho_2 \sigma_i$ and $\gamma_4 = \rho_3$, the matrix γ_5 is

$$\gamma_5 = -\rho_1 \quad . \tag{3.103}$$

When $m = 0$, (3.26) becomes simply

$$\vec{\alpha}\cdot\vec{p}\begin{cases} u_{\vec{p},s} \\ v_{-\vec{p},s} \end{cases} = |\vec{p}|\begin{cases} u_{\vec{p},s} \\ -v_{-\vec{p},s} \end{cases},$$

which, in the representation $\vec{\alpha} = \rho_1\vec{\sigma}$ and $\gamma_5 = -\rho_1$ can also be written as

$$\gamma_5\vec{\sigma}\cdot\hat{p}\begin{cases} u_{\vec{p},s} \\ v_{-\vec{p},s} \end{cases} = \begin{cases} -u_{\vec{p},s} \\ v_{-\vec{p},s} \end{cases}.$$

By comparing this expression with (3.27) we find

$$\begin{aligned} \gamma_5 u_{\vec{p},s} &= -2s\, u_{\vec{p},s} \\ \text{and} \quad \gamma_5 v_{-\vec{p},s} &= 2s\, v_{-\vec{p},s} \end{aligned} \quad . \tag{3.104}$$

Thus for any $\vec{p}$, $u_{\vec{p},-\frac{1}{2}}$ and $v_{\vec{p},\frac{1}{2}}$ are eigenvectors of γ_5 with an eigenvalue $+1$, while $u_{\vec{p},\frac{1}{2}}$ and $v_{\vec{p},-\frac{1}{2}}$ are those of γ_5 with an eigenvalue -1. Because, according to (3.101),

$$\gamma_5\psi_L = \psi_L \quad \text{and} \quad \gamma_5\psi_R = -\psi_R,$$

the Fourier expansion (3.32) can be separated into two parts

$$\psi_L = \frac{1}{\sqrt{\Omega}}\sum_{\vec{p}}\{a_{\vec{p},-\frac{1}{2}}\, u_{\vec{p},-\frac{1}{2}}\, e^{i\vec{p}\cdot\vec{r}} + b^{\dagger}_{\vec{p},+\frac{1}{2}}\, v_{\vec{p},+\frac{1}{2}}\, e^{-i\vec{p}\cdot\vec{r}}\}, \tag{3.105}$$

and

$$\psi_R = \frac{1}{\sqrt{\Omega}}\sum_{\vec{p}}\{a_{\vec{p},\frac{1}{2}}\, u_{\vec{p},\frac{1}{2}}\, e^{i\vec{p}\cdot\vec{r}} + b^{\dagger}_{\vec{p},-\frac{1}{2}}\, v_{\vec{p},-\frac{1}{2}}\, e^{-i\vec{p}\cdot\vec{r}}\}. \tag{3.106}$$

Since the equation $\psi = \psi_L + \psi_R$ is simply a restatement of (3.32), it is therefore valid in all cases, whether $m = 0$ or not. However, this decomposition is a particularly useful one in the case of $m = 0$. The free Hamiltonian density (3.35) is invariant under a phase transformation

$$\psi \to e^{i\theta}\psi. \tag{3.107}$$

When $m = 0$, (3.35) becomes simply

$$\mathcal{H}_{\text{free}} = \psi^{\dagger}(-i\vec{\alpha}\cdot\vec{\nabla})\,\psi$$

which, because $[\gamma_5, \vec{a}] = 0$, is also invariant under the transformation

$$\psi \rightarrow \gamma_5 \psi \quad . \tag{3.108}$$

Therefore, we may impose a supplementary condition: Either $\gamma_5 \psi = \psi$, and as a result

$$\psi = \psi_L \quad , \tag{3.109}$$

or $\gamma_5 \psi = -\psi$, so that

$$\psi = \psi_R \quad . \tag{3.110}$$

Clearly this is possible only when $m = 0$. Physically, this can also be understood in a simple way: When $m = 0$, under a Lorentz transformation the particle momentum $\vec{p} \rightarrow \vec{p}'$, but its helicity s remains unchanged. However, if $m \neq 0$, for any given $\vec{p}$ we can always perform a Lorentz transformation along $\vec{p}$ but with a velocity larger than $\vec{p}/E_p$, whereupon the particle-momentum direction will change sign but its spin direction will be the same. Consequently, the particle helicity s will be changed to $-s$. Such a Lorentz transformation is not possible when $m = 0$. This explains why, in accordance with (3.105) and (3.106), ψ_L can only annihilate particle states with helicity $s = -\frac{1}{2}$ (and create antiparticle states with $s = +\frac{1}{2}$) while ψ_R can only annihilate those with $s = +\frac{1}{2}$ (and create antiparticle states with $s = -\frac{1}{2}$).

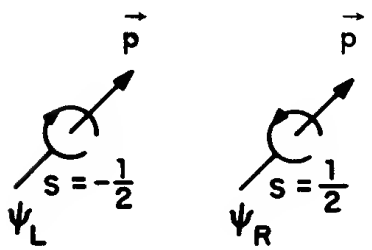


Fig. 3.1. The lefthand field ψ_L can only annihilate particle states with helicity $s = -\frac{1}{2}$, while the righthand field ψ_R can only annihilate those with helicity $s = +\frac{1}{2}$.

Next, we discuss the case with interactions. Let us assume that the interaction Hamiltonian H_{int} is, like the free Hamiltonian, also invariant under the phase transformation (3.107) and the γ_5 transformation (3.108). For example, H_{int} can be a function only of j_μ where

$$j_\mu = i \psi^\dagger \gamma_4 \gamma_\mu (1 + \gamma_5) \psi \quad .$$

In such a case we can impose the supplementary condition (3.109), $\psi = \psi_L$, and require the physical mass of the particle m_{phys} to remain 0. [See Problem 3.2.] All particle states are of helicity $s = -\frac{1}{2}$ and all antiparticle states are of $s = +\frac{1}{2}$. Unlike the non-zero mass case where, in general, for each $\vec{p}$ there are four spinor states, $u_{\vec{p}, \pm\frac{1}{2}}$ and $v_{\vec{p}, \pm\frac{1}{2}}$, here there are only two.

As we shall discuss later, the two-component theory is well-suited for neutrinos.

Remarks. It is possible to transform the matrix representation of $\gamma_1, \gamma_2, \cdots, \gamma_5$ by a unitary transformation

$$\gamma_a \rightarrow U \gamma_a U^\dagger \quad ,$$

where U is a 4×4 unitary matrix. Equations (3.96)-(3.102) and the supplementary condition (3.109), or (3.110), are all invariant under such a transformation. The particular matrix representations (3.1)-(3.11) and (3.103) are adopted for convenience. They lead to equations (3.26), (3.27) and the expansion (3.32), (3.105) and (3.106); in these expansions the subscripts $\vec{p}$ and s refer to the observed momentum and helicity of the particle or antiparticle states.

Problem 3.1. Show that for a free spin-$\frac{1}{2}$ field of mass m

$$\{\psi_a(x), \bar{\psi}_\beta(0)\} = i \left(\gamma_\mu \frac{\partial}{\partial x_\mu} - m\right)_{a\beta} D(x)$$

where $D(x)$ is given in Problem 2.1, and $\bar{\psi} = \psi^\dagger \gamma_4$.

Problem 3.2. Let ψ be a usual quantized four-component Dirac spinor field, and ψ_L be the corresponding two-component field defined by (3.101):

$$\psi_L \equiv \tfrac{1}{2}(1 + \gamma_5)\,\psi \quad .$$

We may define a two-component theory to be one in which the Lagrangian density is a function only of ψ_L, $\psi_L^\dagger$ and possibly also other fields, but not of ψ_R and $\psi_R^\dagger$. Clearly, a two-component theory is always invariant under the γ_5 transformation: $\psi \to \gamma_5 \psi$.

(i) Show that a two-component theory which is also invariant under the phase transformation (3.107), $\psi \to e^{i\theta}\psi$, implies a zero physical mass for the particle. [Such a theory may be called the Weyl theory.]

As we shall discuss later in Chapter 10, the phase invariance is connected with the fermion-number conservation law.

(ii) Assume that the Lagrangian density of a two-component theory is given by

$$\mathcal{L} = -\psi_L^\dagger \gamma_4 \gamma_\mu \frac{\partial}{\partial x_\mu} \psi_L - \frac{m}{2}(\psi_L^\dagger \gamma_4 \psi_L^c + \text{h.c.})$$

where

$$(\psi_L^c)_a = (\gamma_2)_{a\beta} (\psi_L^\dagger)_\beta \quad .$$

The Hamiltonian density is then given by

$$\mathcal{H} = \psi_L^\dagger (-i\vec{\alpha}\cdot\vec{\nabla})\,\psi_L + \frac{m}{2}(\psi_L^\dagger \beta \psi_L^c + \text{h.c.})$$

and h.c. denotes the Hermitian conjugate. Calculate the energy spectrum in this theory and show that the physical mass of the particle is m.

Thus, a two-component theory that is not invariant under the

phase transformation (3.107) can acquire a nonzero mass. [Such a theory may be called the Majorana theory.*]

(iii) Show that identical conclusions can be reached if ψ_L is replaced by $\psi_R = \frac{1}{2}(1 - \gamma_5)\psi$.

Hint for (ii): Define the Majorana field operator

$$\psi_M = (\psi_L + \psi_L^c)/\sqrt{2} \quad ,$$

which is invariant under the particle-antiparticle conjugation; i.e.,

$$\psi_M^c = \psi_M$$

where $(\psi_M^c)_a = (\gamma_2)_{a\beta}(\psi_M^\dagger)_\beta$. Note that the Lagrangian density

$$\mathcal{L}_M \equiv -\psi_M^\dagger \gamma_4 (\gamma_\mu \frac{\partial}{\partial x_\mu} + m)\psi_M$$

differs from $\mathcal{L}$ only by a total derivative.

References

P. A. M. Dirac, Quantum Mechanics (Oxford, The Clarendon Press, 1958).

W. Pauli, Reviews of Modern Physics 13, 203 (1941).

G. Wentzel, Quantum Theory of Fields (New York, Interscience Publishers, Inc., 1949).

---

* See K. Case, Phys.Rev. 107, 307 (1957).

Chapter 4

THE SPIN-1 FIELD ($m \neq 0$)

4.1 Free Field

The Lagrangian density for a free spin-1 field with mass $m \neq 0$ is

$$\mathcal{L} = \mathcal{L}_{\text{free}} = -\tfrac{1}{4} F_{\mu\nu}^2 - \tfrac{1}{2} m^2 A_\mu^2 \quad , \tag{4.1}$$

where, like x_μ, the space components of A_μ are Hermitian and the time component anti-Hermitian, i.e.

$$\vec{A} = \vec{A}^\dagger \quad , \qquad A_4 = -A_4^\dagger \quad ,$$

and

$$F_{\mu\nu} = \frac{\partial}{\partial x_\mu} A_\nu - \frac{\partial}{\partial x_\nu} A_\mu \quad . \tag{4.2}$$

From the variational principle

$$\delta \int \mathcal{L}\, d^3r\, dt = 0 \quad , \tag{4.3}$$

we find

$$\frac{\partial}{\partial x_\mu} F_{\mu\nu} - m^2 A_\nu = 0 \quad . \tag{4.4}$$

By taking its divergence

$$\frac{\partial}{\partial x_\nu} \left(\frac{\partial}{\partial x_\mu} F_{\mu\nu} - m^2 A_\nu \right) = 0$$

and by noting that $m \neq 0$ and $F_{\mu\nu} = -F_{\nu\mu}$, we derive for the free field

$$\frac{\partial}{\partial x_\mu} A_\mu = 0 \quad . \tag{4.5}$$

As will be discussed later, this equation is not necessarily valid when

there are interactions.

To carry out the quantization, we observe that in (4.1) $\mathcal{L}$ does not contain $\dot{A}_4$. Thus we may regard A_4 as a dependent variable. From

$$F_{4j} = -i\dot{A}_j - \nabla_j A_4$$

and

$$-\tfrac{1}{4}F_{\mu\nu}^2 = \tfrac{1}{2}(\dot{A}_j - i\nabla_j A_4)^2 - \tfrac{1}{2}(\vec{\nabla}\times\vec{A})^2 ,$$

we find the conjugate momenta of A_j to be

$$\Pi_j = \frac{\partial\mathcal{L}}{\partial\dot{A}_j} = \dot{A}_j - i\nabla_j A_4 = iF_{4j} , \tag{4.6}$$

where, as before, all Roman subscripts i, j vary from 1 to 3, all Greek subscripts μ, ν vary from 1 to 4, and all repeated indices are to be summed over. By setting $\nu = 4$ in (4.4), we have

$$A_4 = \frac{1}{m^2}\nabla_j F_{j4} = \frac{i}{m^2}\vec{\nabla}\cdot\vec{\Pi} , \tag{4.7}$$

which, together with (4.6), gives

$$\vec{\Pi} = \dot{\vec{A}} + \frac{1}{m^2}\vec{\nabla}(\vec{\nabla}\cdot\vec{\Pi}) . \tag{4.8}$$

In order to derive the Hamiltonian, we regard $\vec{A}$ and $\vec{\Pi}$ as independent variables, but A_4 as a function of $\vec{\Pi}$ through (4.7). From (4.8), it follows that

$$\vec{\Pi}\cdot\dot{\vec{A}} = \vec{\Pi}^2 - \frac{1}{m^2}\vec{\Pi}\cdot\vec{\nabla}(\vec{\nabla}\cdot\vec{\Pi}) ,$$

and therefore

$$\int \vec{\Pi}\cdot\dot{\vec{A}}\, d^3r = \int [\vec{\Pi}^2 + \frac{1}{m^2}(\vec{\nabla}\cdot\vec{\Pi})^2]\, d^3r .$$

Thus

$$H_{free} = \int (\vec{\Pi}\cdot\dot{\vec{A}} - \mathcal{L}_{free})\, d^3r$$

$$= \int \tfrac{1}{2}[\vec{\Pi}^2 + (\frac{\vec{\nabla}\cdot\vec{\Pi}}{m})^2 + (\vec{\nabla}\times\vec{A})^2 + m^2\vec{A}^2]\, d^3r . \tag{4.9}$$

According to the usual quantization rule, the equal-time commutator

between $A_i(\vec{r}, t)$ and $\Pi_j(\vec{r}', t)$ is

$$[A_i(\vec{r}, t), \Pi_j(\vec{r}', t)] = i\delta_{ij}\delta^3(\vec{r} - \vec{r}') \tag{4.10}$$

while all other equal-time commutators $[A_i, A_j]$ and $[\Pi_i, \Pi_j]$ are zero.

To carry out the Fourier expansion, we introduce for any given $\vec{k}$ a set of three orthogonal unit vectors $\hat{k} = \vec{k}/|\vec{k}|$, $\hat{e}_1$ and $\hat{e}_2$:

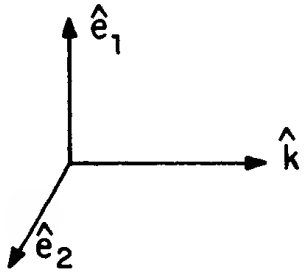


Fig. 4.1. An orthonormal set of three vectors.

At any fixed time t we may expand $\vec{A}(\vec{r}, t)$ and $\vec{\Pi}(\vec{r}, t)$ in terms of the Fourier series:

$$\vec{A}(\vec{r}, t) = \sum_{\vec{k}} \frac{1}{\sqrt{2\omega\,\Omega}} \left\{ [\hat{k}\, a_L(\vec{k}) \frac{\omega}{m} + \sum_{T=1,2} \hat{e}_T\, a_T(\vec{k})]\, e^{i\vec{k}\cdot\vec{r}} + \text{h.c.} \right\} \tag{4.11}$$

and

$$\vec{\Pi}(\vec{r}, t) = \sum_{\vec{k}} \frac{1}{\sqrt{2\omega\,\Omega}} \left\{ -i[m\hat{k}\, a_L(\vec{k}) + \omega \sum_{T=1,2} \hat{e}_T\, a_T(\vec{k})]\, e^{i\vec{k}\cdot\vec{r}} + \text{h.c.} \right\} \tag{4.12}$$

in which h.c. denotes the Hermitian conjugate terms and, as before in (2.17),

$$\omega = \sqrt{\vec{k}^2 + m^2} \quad . \tag{4.13}$$

The subscripts L and T indicate the longitudinal and transverse components respectively and the $a_L(\vec{k})$'s and $a_T(\vec{k})$'s are all functions of t. From (4.10) and by following similar arguments used in

Chapter 2 to derive (2.25), we obtain the following equal-time commutator between $a_i(\vec{k})$ and $a_j^\dagger(\vec{k}')$:

$$[a_i(\vec{k}),\ a_j^\dagger(\vec{k}')] = \delta_{ij}\,\delta_{\vec{k},\vec{k}'} \quad . \tag{4.14}$$

All other equal-time commutators are zero; i.e.

$$[a_i(\vec{k}),\ a_j(\vec{k}')] = 0 \quad \text{and} \quad [a_i^\dagger(\vec{k}),\ a_j^\dagger(\vec{k}')] = 0 \tag{4.15}$$

where i and j denote either the longitudinal component L or the transverse components T = 1 or 2 .[See the exercise on the next page.] In terms of these annihilation and creation operators, H_{free} is

$$H_{free} = \sum_{\vec{k}} \omega [a_L^\dagger(\vec{k})\, a_L(\vec{k}) + \sum_{T=1,2} a_T^\dagger(\vec{k})\, a_T(\vec{k}) + \tfrac{3}{2}] \ . \tag{4.16}$$

By using Heisenberg's equation we see that

$$a_i(\vec{k}) \propto e^{-i\omega t} \quad \text{and} \quad a_i^\dagger(\vec{k}) \propto e^{i\omega t} \quad . \tag{4.17}$$

From (4.7) and (4.12), it follows that

$$A_4(\vec{r}, t) = i \sum_{\vec{k}} \frac{1}{\sqrt{2\omega\,\Omega}} \frac{|\vec{k}|}{m} a_L(\vec{k})\, e^{i\vec{k}\cdot\vec{r}} + \text{h.c.} \tag{4.18}$$

By substituting (4.11), (4,17) and (4.18) into (4.5), we can verify that, for free fields, $\partial A_\mu / \partial x_\mu = 0$.

4.2 Interacting Fields

Next we consider the case that the spin-1 field A_μ has interactions with other fields. For simplicity, let us assume the Lagrangian density to be given by

$$\mathcal{L} = -\tfrac{1}{4} F_{\mu\nu}^2 - \tfrac{1}{2} m^2 A_\mu^2 - j_\mu A_\mu + \cdots , \tag{4.19}$$

where j_μ and the $\cdots$ terms depend only on other fields, independent of A_μ .

From the action principle (4.3), we have

$$\frac{\partial}{\partial x_\mu} F_{\mu\nu} - m^2 A_\nu = j_\nu \quad , \tag{4.20}$$

from which it follows thot

$$\frac{\partial}{\partial x_\mu} A_\mu = - \frac{1}{m^2} \frac{\partial}{\partial x_\nu} j_\nu \quad . \tag{4.21}$$

Since, by assumption, the current j_ν is independent of A_μ, the question whether j_ν is conserved or not depends on the equotions of motion of other fields and on the detoiled form of the current. If $\partial j_\nu / \partial x_\nu = 0$, then $\partial A_\mu / \partial x_\mu = 0$; otherwise not.

Just os in the previous section, the new Logrongian (4.19) does not contain $\dot{A}_4$; we shall regord A_4 as a dependent voriable. From (4.19) we see thot the conjugote momentum Π_j remoins given by (4.6); i.e.

$$\Pi_j = \frac{\partial \mathcal{L}}{\partial \dot{A}_j} = i F_{4j} \quad . \tag{4.22}$$

By setting $\nu = 4$ in (4.20), we find that (4.7) should be replaced by

$$A_4 = \frac{i}{m^2} \vec{\nabla} \cdot \vec{\Pi} - \frac{1}{m^2} j_4 \quad . \tag{4.23}$$

The equol-time commutotor between A_i ond Π_j is still given by (4.10) and, as before, oll other equal-time commutators $[A_i, A_j]$ ond $[\Pi_i, \Pi_j]$ remain zero.

Exercise. In the present cose of interocting fields, show that ot ony time t we may still expand $\vec{A}(\vec{r}, t)$ and $\vec{\Pi}(\vec{r}, t)$ in terms of the Fourier series (4.11)-(4.12). In oddition, the equol-time commutation relotions (4.14)-(4.15) remoin volid.

From the Lagrangian density (4.19) ond by regording A_4 as o function of $\vec{\Pi}$ ond other field voriables through (4.23), the Homiltonian density con be constructed in the usuol woy. Unlike the free-field

case, the time dependences of $a_i(\vec{k})$ and $a_i^\dagger(\vec{k})$ are in general quite complicated, in contrast to (4.17).

Remarks. The presence of A_4 makes some of the discussions of a spin-1 field quite different from those of a spin-0 field. For a given momentum $\vec{k}$, a spin-1 particle with a non-zero physical mass, $m \neq 0$, has three modes: longitudinal L and transverse T = 1, 2. They correspond to, in the rest frame of the particle, the three z-component angular momentum states $j_z = 0$ and ± 1. On the other hand, because of Lorentz invariance, the spin-1 field is represented by A_μ which has four components. Consequently, one of the four components is not an independent variable.

Because in many of the above expressions the parameter m is often in the denominator, the limit $m = 0$ for a spin-1 field is more complicated than that for a spin-$\frac{1}{2}$ field. The details will be given in Chapter 6, when we discuss quantum electrodynamics.

Problem 4.1. Consider a system which consists of a spin-1 field A_μ and a spin-$\frac{1}{2}$ field ψ. Assume that the Lagrangian density is given by (4.19) in which the ··· term is

$$-\psi^\dagger \gamma_4 (\gamma_\mu \frac{\partial}{\partial x_\mu} + m') \psi$$

and j_μ is either

(i) $j_\mu = i g \psi^\dagger \gamma_4 \gamma_\mu \psi$

or

(ii) $j_\mu = i g \psi^\dagger \gamma_4 \gamma_\mu \gamma_5 \psi$.

Construct the Hamiltonian for this system and work out the details of the Fourier expansion for ψ, $\vec{A}$ and A_4. Use Heisenberg's equation to obtain the equation of motion. Find out in which of the above

cases $\partial A_\mu / \partial x_\mu = 0$.

Problem 4.2. Show that for a free spin-1 field of mass m

$$[A_\mu(x), A_\nu(0)] = -i\left(\delta_{\mu\nu} - m^{-2}\frac{\partial^2}{\partial x_\mu \partial x_\nu}\right) D(x)$$

where $D(x)$ is given in Problem 2.1.

References

W. Pauli, Reviews of Modern Physics 13, 203 (1941).

G. Wentzel, Quantum Theory of Fields (New York, Interscience Publishers, Inc., 1949).

Chapter 5

FEYNMAN DIAGRAMS

5.1 Heisenberg, Schrödinger and Interaction Representations

In quantum mechanics all experimental results can be expressed in terms of the matrix elements $< a \mid O \mid b >$ of different operators O between various state vectors $\mid a >$ and $\mid b >$. There are many ways to describe the time variation of such matrix elements.

1. Heisenberg representation

In the Heisenberg representation only the physical operators vary with time. All physical state vectors are time-independent. If we denote $O_H(t)$ and $\mid t >_H$ as the operator $O(t)$ and the state vector $\mid t >$ in the Heisenberg representation, their equations of motion are

$$[H_H , O_H(t)] = - i \dot{O}_H(t) \tag{5.1}$$

and

$$\frac{\partial}{\partial t} \mid t >_H = 0 , \tag{5.2}$$

where H_H is the Hamiltonian in the Heisenberg representation.

2. Schrödinger representation

In the Schrödinger representation only the physical state vectors are time-dependent; all physical operators are time-independent. Let $O_S(t)$ and $\mid t >_S$ be the operator $O(t)$ and the state vector $\mid t >$ in the Schrödinger representation. The equations of motion are

$$\dot{O}_S(t) = 0 \tag{5.3}$$

and

$$H_S \mid t >_S = -\frac{1}{i}\frac{\partial}{\partial t} \mid t >_S \ , \tag{5.4}$$

where H_S is the Hamiltonian in the Schrödinger representation.

3. Interaction representation

Let H_I be the Hamiltonian in the interaction representation, which will be decomposed into

$$H_I = (H_0)_I + (H_{int})_I \ . \tag{5.5}$$

We denote $O_I(t)$ and $\mid t >_I$ as the operator $O(t)$ and the state vector $\mid t >$ in the interaction representation. The equations of motion are now

$$[\,(H_0)_I\ ,\ O_I(t)\,] = -i\,\dot{O}_I(t) \ , \tag{5.6}$$

and

$$(H_{int}(t))_I \mid t >_I = -\frac{1}{i}\frac{\partial}{\partial t} \mid t >_I \ . \tag{5.7}$$

So far, the decomposition (5.5) is quite arbitrary. If $H_0 = 0$, then the interaction representation is identical to the Schrödinger representation; if $H_{int} = 0$, then it is identical to the Heisenberg representation.

We shall now show the equivalence of these different representations. For clarity of notation, we shall write

$$H = H_S \ , \tag{5.8}$$

i.e., we shall omit the subscript S when the Hamiltonian is in the Schrödinger representation. Likewise, the decomposition (5.5) will be written in the Schrödinger representation as

$$H = H_0 + H_{int} \ ,$$

i.e., just as in (5.8),

$$H_0 = (H_0)_S \qquad \text{and} \qquad H_{int} = (H_{int})_S \ . \tag{5.9}$$

In the Schrödinger representation, the state $| t >_S$ can be readily expressed in terms of its form $| 0 >_S$ at $t = 0$:

$$| t >_S = e^{-iHt} | 0 >_S \quad . \tag{5.10}$$

The corresponding state vector $| t >_H$ in the Heisenberg representation is time-independent, and may be set to be equal to that in the Schrödinger representation at $t = 0$; i.e.

$$| t >_H = | 0 >_H = | 0 >_S \qquad \text{at all } t \ .$$

Hence, the state vectors $| t >_H$ and $| t >_S$ are related by the unitary transformation e^{iHt}:

$$| t >_H = e^{iHt} | t >_S \quad . \tag{5.11}$$

Under the same unitary transformation, the corresponding operators in these two representations are related by

$$O_H(t) = e^{iHt} O_S e^{-iHt} \quad . \tag{5.12}$$

By setting the operator $O(t)$ to be the total Hamiltonian, we see that it is unchanged when we switch from the Schrödinger representation to the Heisenberg representation. Hence, in the notation (5.8),

$$H_H = H_S = H \quad . \tag{5.13}$$

By differentiating (5.11) and (5.12), and by using (5.3)–(5.4) and (5.13) we see that

$$\begin{aligned} -i \dot{O}_H(t) &= -i \frac{\partial}{\partial t} (e^{iHt} O_S e^{-iHt}) \\ &= (H e^{iHt} O_S e^{-iHt} - e^{iHt} O_S e^{-iHt} H) \\ &= H O_H(t) - O_H(t) H = [H_H, O_H] \end{aligned}$$

and

$$-\frac{1}{i} \frac{\partial}{\partial t} | t >_H = e^{iHt}(-H + H) | t >_S = 0 \quad .$$

Hence the equations of motion (5.3)-(5.4) in the Schrödinger representation imply those in the Heisenberg representation. Similarly, we can verify that the converse is also correct.

The state vector $| t >_I$ in the interaction representation is related to that in the Schrödinger representation by

$$| t >_I = e^{iH_0 t} | t >_S \tag{5.14}$$

where H_0 is given by (5.9). Under the same unitary transformation $e^{iH_0 t}$, the corresponding operators in these two representations are related by

$$O_I(t) = e^{iH_0 t} O_S e^{-iH_0 t} . \tag{5.15}$$

By setting the operator $O(t)$ to be H_0 we see that

$$(H_0)_I = (H_0)_S = H_0 . \tag{5.16}$$

Hence, H_0 is unchanged when we switch from the Schrödinger representation to the interaction representation. In the following, H_0 will be called the unperturbed Hamiltonian. By differentiating (5.14) and (5.15) and by using (5.3)-(5.4) and (5.9), we find

$$\begin{aligned} -i \dot{O}_I(t) &= -i \frac{\partial}{\partial t} (e^{iH_0 t} O_S e^{-iH_0 t}) \\ &= H_0 e^{iH_0 t} O_S e^{-iH_0 t} - e^{iH_0 t} O_S e^{-iH_0 t} H_0 \\ &= [(H_0)_I , O_I(t)] \end{aligned}$$

and

$$\begin{aligned} -\frac{1}{i} \frac{\partial}{\partial t} | t >_I &= -\frac{1}{i} \frac{\partial}{\partial t} e^{iH_0 t} | t >_S \\ &= e^{iH_0 t}(-H_0 + H) | t >_S \\ &= e^{iH_0 t} H_{int} | t >_S \end{aligned}$$

$$= e^{iH_0 t} H_{int} e^{-iH_0 t} e^{iH_0 t} | t >_S$$

$$= (H_{int}(t))_I | t >_I \ .$$

Therefore, the equations of motion (5.3)-(5.4) in the Schrödinger representation also imply those in the interaction representation. Likewise we can show that the converse is also correct. From (5.11)-(5.12) and (5.14)-(5.15), it follows that at any time t

$$(< a | O | b >)_H = (< a | O | b >)_S = (< a | O | b >)_I$$

where O can be any operator and $| a >$, $| b >$ can be any two state vectors. That completes the proof of equivalence between these different representations. [See also (24.100).]

5.2 S - Matrix

In this section, as well as throughout the remainder of this chapter, we shall stay in the interaction representation. For simplicity, the subscript I will be omitted. Therefore (5.7) becomes

$$-\frac{1}{i}\frac{\partial}{\partial t} | t > = H_{int}(t) | t > \qquad (5.17)$$

where in accordance with (5.6) $H_{int}(t)$ satisfies

$$-i\dot{H}_{int} = [H_0 , H_{int}] \ . \qquad (5.18)$$

Both H_0 and H_{int} are Hermitian. Let $U(t, t_0)$ be the Green's function of (5.17). Then we have

$$| t > = U(t, t_0) | t_0 > \qquad (5.19)$$

where $U(t, t_0)$ satisfies

$$-\frac{1}{i}\frac{\partial}{\partial t} U(t, t_0) = H_{int}(t) U(t, t_0) \ , \qquad (5.20)$$

with the initial condition

$$U(t_0, t_0) = 1 = \text{unit matrix} . \tag{5.21}$$

The Hermitian conjugate of (5.20) is

$$\frac{1}{i} \frac{\partial}{\partial t} U^\dagger(t, t_0) = U^\dagger(t, t_0) H_{int}(t) .$$

Thus,

$$\frac{1}{i} \frac{\partial}{\partial t} (U^\dagger(t, t_0) U(t, t_0)) = 0$$

which, together with the initial condition (5.21), implies

$$U^\dagger(t, t_0) U(t, t_0) = 1 , \tag{5.22}$$

i.e., $U(t, t_0)$ is unitary.

The S-matrix is defined to be the limit

$$S \equiv \lim_{\substack{t_0 \to -\infty \\ t \to \infty}} U(t, t_0) . \tag{5.23}$$

It connects the state vector from time $= -\infty$ to $+\infty$. Any scattering problem can be described by the transformation between the initial and final state vectors; the former is given in the remote past ($t_0 = -\infty$) while the latter is in the remote future ($t = +\infty$). The scattering amplitude is therefore given by the corresponding matrix element of the S-matrix.

So far the decomposition $H = H_0 + H_{int}$ is quite arbitrary. However, in order for the double limits (5.23) to exist, there are some simple requirements.

To see this, let us consider a simple example of a single harmonic oscillator whose Hamiltonian is given by

$$H = \omega a^\dagger a$$

where a and $a^\dagger$ are the usual annihilation and creation operators. Suppose H is decomposed into the sum of

$$H_0 = \omega_0 a^\dagger a ,$$

and

$$H_{int} = (\omega - \omega_0)\, a^\dagger a \quad . \tag{5.24}$$

The solution of (5.20) is

$$U(t, t_0) = e^{-i(\omega-\omega_0)(t-t_0)\, N} \tag{5.25}$$

where N is the occupation number operator

$$N = a^\dagger a \quad .$$

Let $| n >$ be the eigenvector of N with eigenvalue n. The diagonal matrix element $< n \mid U(t, t_0) \mid n >$ is

$$e^{-i(\omega-\omega_0)(t-t_0)\, n} \quad ;$$

for any $n \neq 0$; its limit when either $t_0 \to -\infty$ or $t \to \infty$ does not exist unless $\omega = \omega_0$.

There are several different approaches by which we can bypass this difficulty. One way is to modify the differential equation (5.18) or (5.20) in the asymptotic region in time. For example, one may keep H_{int} intact but assume t to have an appropriate small imaginary component, or in a completely equivalent way, keep t real but replace (5.24) by

$$H_{int}(t) = \begin{cases} (\omega-\omega_0)\, a^\dagger a & \text{for } | t | \leqslant T \\ (\omega-\omega_0 - i\epsilon)\, a^\dagger a & \text{for } | t | > T > 0 \end{cases}$$

where $\epsilon = 0+$ and T is a large constant. Thus, (5.25) is valid only when t_0 and t are both within the range $-T$ to T, but not outside; e.g. when $t_0 < -T$ and $t > T$, (5.25) is replaced by

$$U(t, t_0) = e^{-i(\omega-\omega_0)(t-t_0)\, N}\, e^{-\epsilon(t-t_0-2T)\, N}$$

which, at a fixed $\epsilon > 0$, does have a limit as $t \to \infty$ and $t_0 \to -\infty$, but this limit is not unitary.

Of course, one may try a different modification: For example, (5.24) may be replaced by

$$H_{int}(t) = \begin{cases} (\omega - \omega_0)\, a^\dagger a & \text{for } |t| \leqslant T \\ 0 & \text{for } |t| > T \end{cases} .$$

As above, (5.25) remains valid when t_0 and t are within the range $-T$ to T; outside this range, for $t > T$ and $t_0 < -T$, we have

$$U(t, t_0) = U(T, -T) = e^{-i\,2(\omega-\omega_0)\,TN} .$$

Thus, the limit (5.23) exists and is now unitary; however, it depends on an artificial parameter T.

There exists still another method, and this is the one we shall adopt. There is no change in either of the differential equations (5.18) or (5.20) and t remains real. However, in this approach we require the spectrum of H_0 to be identical to that of H. As we shall see, for a large class of relativistic theories, it turns out that this seemingly stringent condition can be easily satisfied. In the above simple example, this condition leads to $\omega = \omega_0$ and therefore $H_{int} = 0$ and $U = S = 1$.

Next, let us consider a spin-0 field ϕ whose Hamiltonian density is given by

$$H = \tfrac{1}{2}(\Pi^2 + (\vec{\nabla}\phi)^2 + m_0^2\,\phi^2) + g_0^2\,\phi^4 - E_{vac} , \qquad (5.26)$$

where Π is the conjugate momentum of ϕ, and m_0, g_0 and E_{vac} are real numbers. The parameter m_0 is called the mechanical mass and g_0 the unrenormalized coupling constant. We define the vacuum state $|\,vac>$ to be the lowest-energy state of H. The parameter E_{vac} in (5.26) is determined by the condition

$$H\,|\,vac> = 0 .$$

Since in (5.26) the interaction $g_0^2\,\phi^4$ is positive, it gives rise only to repulsive forces between the spin-0 quanta. We may therefore

assume that there is no bound state in the system. From relativistic invariance we know that the energy spectrum of the system must be of the form

$$E = \sum_{\vec{k}} n_{\vec{k}} \sqrt{\vec{k}^2 + m^2} \tag{5.27}$$

where m is the physical mass and $n_{\vec{k}} = 0, 1, 2, \cdots$, representing the number of physical spin-0 quanta with momentum $\vec{k}$. Consequently, we may introduce

$$H_0 = \tfrac{1}{2}(\Pi^2 + (\vec{\nabla}\phi)^2 + m^2 \phi^2) - E_0 \ , \tag{5.28}$$

and E_0 is chosen so that the lowest energy state $|0>$ of H_0 also has a zero eigenvalue, i.e.

$$H_0 |0> = 0 \ . \tag{5.29}$$

Clearly, the spectrum of H_0 is identical to that of H provided the parameter m in (5.28) is the physical mass.

In a general case of interacting relativistic fields, when the theory has no stable bound states (as in the above example or in quantum electrodynamics, to be discussed in the next chapter), the unperturbed Hamiltonian H_0 will be chosen to be that of free fields, but with their mass parameters set to be the physical masses of the interacting system. Consequently, the spectrum of H_0 becomes the same as that of H, which makes it possible to take the double limits (5.23) for the S-matrix.

In the case that there are stable bound states, we shall still adopt the same procedure, setting H_0 to be the free-particle Hamiltonian with the m's the physical mass parameters. This enables us to make the formal perturbation series expansion in terms of diagrams, as will be discussed in the subsequent sections. Each stable bound state can be identified as an infinite sum of an appropriate set of

diagrams. After isolating these infinite sums, we can then carry out the limit (5.23). For the moment, we shall ignore such complications.

Therefore, in our interaction representation all field operators satisfy the free equations; their time-dependences are known. For example, the Fourier transformation (2.20) can then be written as

$$\phi(x) = a(x) + a^{\dagger}(x) \quad , \tag{5.30}$$

where

$$a(x) = \sum_{\vec{k}} \frac{1}{\sqrt{2\omega\,\Omega}}\, a_{\vec{k}}\, e^{i\vec{k}\cdot\vec{r} - i\omega t} \quad , \tag{5.31}$$

$a^{\dagger}(x)$ is its Hermitian conjugate, as before $x = (\vec{r}, it)$ and

$$\omega = \sqrt{\vec{k}^2 + m^2} \; > \; 0$$

in which m is chosen to be the physical mass of the spin-0 field. Here, the operator $a_{\vec{k}}$ is time-independent, whereas the operator $a_{\vec{k}}(t)$ in (2.20) has a time-dependence $e^{-i\omega t}$ in the interaction representation; these two are related by $a_{\vec{k}}(t) = a_{\vec{k}}\, e^{-i\omega t}$.

Likewise, for a spin-$\frac{1}{2}$ field the Fourier expansion (3.32) in the interaction representation can be written as

$$\psi(x) = u(x) + \bar{v}(x) \quad , \tag{5.32}$$

where

$$u(x) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{p},s} a_{\vec{p},s}\, u_{\vec{p},s}\, e^{i\vec{p}\cdot\vec{r} - iE_p t} \quad , \tag{5.33}$$

$$\bar{v}(x) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{p},s} b^{\dagger}_{\vec{p},s}\, v_{\vec{p},s}\, e^{-i\vec{p}\cdot\vec{r} + iE_p t} \quad , \tag{5.34}$$

and

$$E_p = \sqrt{\vec{p}^2 + m^2} \; > \; 0 \quad ,$$

with m the physical mass of the spin-$\frac{1}{2}$ field. Here, the operators $a_{\vec{p},s}$ and $b^{\dagger}_{\vec{p},s}$ are again both time-independent; they are related to the corresponding operators $a_{\vec{p},s}(t)$ and $b^{\dagger}_{\vec{p},s}(t)$ in (3.32) by $a_{\vec{p},s}(t) = a_{\vec{p},s}\, e^{-iE_p t}$ and $b^{\dagger}_{\vec{p},s}(t) = b^{\dagger}_{\vec{p},s}\, e^{iE_p t}$. It is convenient to

introduce

$$\bar{\psi} \equiv \psi^{\dagger}(x)\,\gamma_4 \equiv \bar{u}(x) + v(x) \quad , \tag{5.35}$$

where

$$\bar{u}(x) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{p},s} a^{\dagger}_{\vec{p},s}\, u^{\dagger}_{\vec{p},s}\, \gamma_4\, e^{-i\vec{p}\cdot\vec{r}+iE_p t} \tag{5.36}$$

and

$$v(x) = \frac{1}{\sqrt{\Omega}} \sum_{\vec{p},s} b_{\vec{p},s}\, v^{\dagger}_{\vec{p},s}\, \gamma_4\, e^{i\vec{p}\cdot\vec{r}-iE_p t} \quad . \tag{5.37}$$

Similar expansions can also be written for a spin-1 field. The details will, however, be omitted here.

5.3 Time-ordered Products, Normal Products and Contractions

In (5.30), (5.32) and (5.35), $a(x)$, $u(x)$ and $v(x)$ contain only annihilation operators, while $a^{\dagger}(x)$, $\bar{u}(x)$ and $\bar{v}(x)$ contain only creation operators. Let $X_i(x_i)$ be any one of these operators at $x_i = (\vec{r}_i\,, it_i)$. We define the time-ordered product of $\prod_{i=1}^{n} X_i(x_i)$ to be

$$T(X_1(x_1)\, X_2(x_2) \cdots X_n(x_n)) \equiv \delta_p X_{p_1}(x_{p_1})\, X_{p_2}(x_{p_2}) \cdots X_{p_n}(x_{p_n}) \,, \tag{5.38}$$

so that the time sequence x_{p_i} satisfies

$$t_{p_1} \geqslant t_{p_2} \geqslant t_{p_3} \geqslant \cdots \geqslant t_{p_n} \quad .$$

Whenever $t_i = t_j$, the relative order of the corresponding operators X_i and X_j is the same on both sides of (5.38). The normal product of $\prod_{i=1}^{n} X_i(x_i)$ is defined to be

$$:X_1(x_1)\, X_2(x_2) \cdots X_n(x_n): \equiv \delta_p X_{p_1}(x_{p_1})\, X_{p_2}(x_{p_2}) \cdots X_{p_n}(x_{p_n}) \,, \tag{5.39}$$

where the permuted sequence on the righthand side is arranged so that annihilation operators always appear to the right of creation operators. In both definitions (5.38) and (5.39), δ_p can be $+1$ or -1,

depending only on the permutation of the fermion operators. If the order of fermion operators on the lefthand side is an even permutation of those on the righthand side then $\delta_p = +1$, otherwise -1. The subscripts $p_1, p_2, \cdots, p_n$ represent a permutation

$$p = \begin{pmatrix} 1 & 2 \cdots n \\ p_1 & p_2 \cdots p_n \end{pmatrix} \quad .$$

The time-ordered product and the normal product of a sum are defined to be the sum of the corresponding products; i.e.

$$T(A+B) = T(A) + T(B)$$

and

$$:A+B: = :A: + :B: \quad . \tag{5.40}$$

The following are a few examples of such products

$$T(\psi_\alpha(1)\ \psi_\beta(2)) = \begin{cases} \psi_\alpha(1)\ \psi_\beta(2) & t_1 \geqslant t_2 \\ -\psi_\beta(2)\ \psi_\alpha(1) & t_1 < t_2 \end{cases} , \tag{5.41}$$

$$T(\phi(1)\ \phi(2)) = \begin{cases} \phi(1)\ \phi(2) & t_1 \geqslant t_2 \\ \phi(2)\ \phi(1) & t_1 < t_2 \end{cases} , \tag{5.42}$$

$$:\psi_\alpha(1)\ \bar{\psi}_\beta(2): = -\bar{u}_\beta(2)\, u_\alpha(1) + u_\alpha(1)\, v_\beta(2) + \bar{v}_\alpha(1)\, v_\beta(2) + \bar{v}_\alpha(1)\, \bar{u}_\beta(2) \quad , \tag{5.43}$$

$$:\phi(1)\ \phi(2): = a(1)\,a(2) + a^\dagger(2)\,a(1) + a^\dagger(1)\,a(2) + a^\dagger(1)\,a^\dagger(2) \tag{5.44}$$

where 1 and 2 stand for x_1 and x_2 respectively, and the subscripts α and β denote the spinor indices.

The Dyson-Wick contraction between $X_1(1)$ and $X_2(2)$ is defined to be

$$\underbrace{X_1(1)\ X_2(2)} \equiv T(X_1(1)\ X_2(2)) - :X_1(1)\ X_2(2): \quad . \tag{5.45}$$

As an example, we may obtain the contraction between $\phi(x)$ and $\phi(0)$ by using (5.42) and (5.44)-(5.45):

$$\underbrace{\phi(x)\,\phi(0)} = \begin{cases} a(x)\,a^{\dagger}(0) - a^{\dagger}(0)\,a(x) & t \geqslant 0 \\ a(0)\,a^{\dagger}(x) - a^{\dagger}(x)\,a(0) & t < 0 \end{cases}$$

where $x = (\vec{r}, it)$. The righthand side of the above expression can be written as commutators. We have

$$\underbrace{\phi(x)\,\phi(0)} = \begin{cases} [a(x),\ a^{\dagger}(0)] & t \geqslant 0 \\ [a(0),\ a^{\dagger}(x)] & t < 0 \end{cases} . \tag{5.46}$$

Likewise, the contraction between $\psi(x)$ and $\bar{\psi}(0)$ is

$$\underbrace{\psi_{\alpha}(x)\,\bar{\psi}_{\beta}(0)} = \begin{cases} \{u_{\alpha}(x),\ \bar{u}_{\beta}(0)\} & t \geqslant 0 \\ -\{v_{\beta}(0),\ \bar{v}_{\alpha}(x)\} & t < 0 \end{cases} . \tag{5.47}$$

We note that these contractions are all c. numbers, while the corresponding time-ordered products and normal products are operators.

Next, we discuss the four-dimensional integral representations of these contractions. By using (5.31) and the commutation relation $[a_{\vec{k}},\ a^{\dagger}_{\vec{k}'}] = \delta_{\vec{k},\vec{k}'}$, we obtain

$$[a(x),\ a^{\dagger}(0)] = \sum_{\vec{k}} \frac{1}{2\omega\Omega}\, e^{i\vec{k}\cdot\vec{r} - i\omega t} ,$$

$$[a(0),\ a^{\dagger}(x)] = \sum_{\vec{k}} \frac{1}{2\omega\Omega}\, e^{-i\vec{k}\cdot\vec{r} + i\omega t} .$$

Let us define the Feynman propagator $D_F(x)$ to be the contraction between $\phi(x)$ and $\phi(0)$. We have

$$D_F(x) = \underbrace{\phi(x)\,\phi(0)} = \sum_{\vec{k}} \frac{1}{2\omega\Omega}\, e^{i\vec{k}\cdot\vec{r} \mp i\omega t} \tag{5.48}$$

in which the minus sign in the exponent holds when t is > 0, and the positive sign when $t < 0$. At $t = 0$ the function $D_F(x)$ is

continuous. We shall now establish the following integral representation

$$D_F(x) \equiv - \frac{i}{(2\pi)^4} \int \frac{e^{ik\cdot x}\, d^4k}{k^2 + (m - i\epsilon)^2} \quad , \tag{5.49}$$

where

$$k \cdot x = \vec{k}\cdot\vec{r} - k_0 t \quad , \quad d^4k = d^3k\, dk_0 \quad ,$$
$$k^2 = \vec{k}^2 - k_0^2 \quad \text{and} \quad k_4 = ik_0 \quad .$$

The parameter ϵ is a positive real infinitesimal; i.e., $\epsilon = 0+$. To see this, we may first perform the integration from $k_0 = -\infty$ to ∞:

$$\int \frac{e^{-ik_0 t}\, dk_0}{\vec{k}^2 - k_0^2 + m^2 - i\epsilon} = \int \frac{-e^{-ik_0 t}\, dk_0}{(k_0 + \omega - i\epsilon)(k_0 - \omega + i\epsilon)} \quad , \tag{5.50}$$

where $\omega = \sqrt{\vec{k}^2 + m^2} > 0$.

In the complex k_0-plane, the integrand of (5.50) has two singularities $k_0 = \omega - i\epsilon$ and $-\omega + i\epsilon$. When $t > 0$, we may consider the contour integral of the same integrand along the closed curve shown in Fig. 5.1(a). When the radius of the half-circle becomes

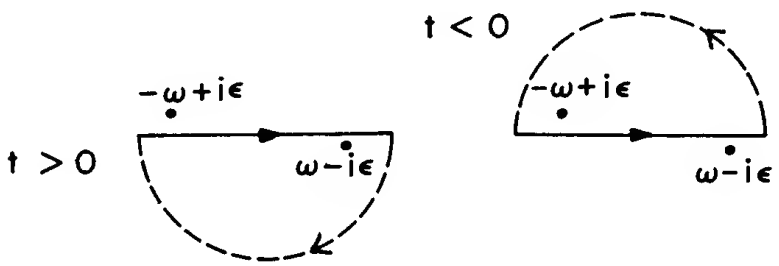


Fig. 5.1. Contours for the Feynman integral (5.46) in the complex k_0-plane.

infinite, the contour integration along this half-circle approaches 0. The integral (5.50) equals $-2\pi i$ times the residue of the integrand at the pole $k_0 = \omega - i\epsilon$. Upon substituting it into (5.49), we obtain for $t > 0$,

$$D_F(x) = \frac{1}{(2\pi)^3} \int \frac{d^3k}{2\omega} e^{i\vec{k}\cdot\vec{r} - i\omega t} \equiv D_+(x) \quad . \tag{5.51}$$

Likewise, when t is negative we consider the contour in Fig. 5.1(b). That leads to, for $t < 0$

$$D_F(x) = \frac{1}{(2\pi)^3} \int \frac{d^3k}{2\omega} e^{i\vec{k}\cdot\vec{r} + i\omega t} \equiv D_-(x) \quad . \tag{5.52}$$

From (2.28) and (5.51)-(5.52), we see the equivalence between (5.48) and (5.49). The $D_+(x)$ denotes the positive frequency part of $D_F(x)$ and the $D_-(x)$ the corresponding negative frequency part.

By using Problem 2.1 and (5.51)-(5.52) we find the commutator between $\phi(x)$ and $\phi(0)$ in the interaction representation to be

$$[\phi(x), \phi(0)] = -iD(x) = D_+(x) - D_-(x) \tag{5.53}$$

where

$$D(x) \equiv \int \frac{1}{8\pi^3} d^3k \frac{\sin \omega t}{\omega} e^{i\vec{k}\cdot\vec{r}} \quad . \tag{5.54}$$

Thus, $D(x)$ and the Feynman propagator $D_F(x)$ are closely related. A more direct way of understanding this relation is to use the decomposition (5.30) and to write

$$[\phi(x), \phi(0)] = [a(x), a^\dagger(0)] + [a^\dagger(x), a(0)] \quad . \tag{5.55}$$

By comparing it with (5.46) we see that its positive frequency part is the same as that in the contraction $D_F(x)$; its negative frequency part differs from the corresponding part in $D_F(x)$ only by a minus sign.

Identical reasoning can be applied to the spin-$\frac{1}{2}$ field ψ. On

account of (5.32) and (5.35), the anticommutator between ψ and its adjoint $\bar{\psi} \equiv \psi^\dagger \gamma_4$ in the interaction representation is

$$\{\psi_\alpha(x),\ \bar{\psi}_\beta(0)\} = \{u_\alpha(x),\ \bar{u}_\beta(0)\} + \{\bar{v}_\alpha(x),\ v_\beta(0)\} \quad (5.56)$$

where, as before, the subscripts α and β are spinor indices. From (5.33) we see that $\{u_\alpha(x), \bar{u}_\beta(0)\}$ consists only of positive frequency terms; i.e. terms proportional to $e^{-iE_p t}$. Likewise from (5.34), $\{\bar{v}_\alpha(x), v_\beta(0)\}$ consists only of negative frequency parts which are proportional to $e^{iE_p t}$. On the other hand, from Problem 3.1 and Eq. (5.53), we see that the same anticommutator is also given by

$$\begin{aligned}\{\psi(x), \bar{\psi}(0)\} &= i(\gamma_\mu \frac{\partial}{\partial x_\mu} - m) D(x) \\ &= -(\gamma_\mu \frac{\partial}{\partial x_\mu} - m)[D_+(x) - D_-(x)] \quad . \quad (5.57)\end{aligned}$$

We now equate, respectively, the positive and negative frequency parts of (5.56) and (5.57). That gives

$$\{u_\alpha(x), \bar{u}_\beta(0)\} = -(\gamma_\mu \frac{\partial}{\partial x_\mu} - m)_{\alpha\beta} D_+(x)$$

and

$$\{\bar{v}_\alpha(x), v_\beta(0)\} = (\gamma_\mu \frac{\partial}{\partial x_\mu} - m)_{\alpha\beta} D_-(x) \quad .$$

By substituting the above expressions into (5.47) and by using (5.51)-(5.52), we derive

$$\underbrace{\psi(x)\ \bar{\psi}(0)} = -(\gamma_\mu \frac{\partial}{\partial x_\mu} - m) D_F(x) \quad (5.58)$$

which, because of (5.49), can also be written as

$$\underbrace{\psi(x)\ \bar{\psi}(0)} = i \int \frac{1}{(2\pi)^4} \frac{i\gamma_\mu k_\mu - m}{k^2 + (m - i\epsilon)^2} e^{ik\cdot x} d^4k \quad .$$

It is convenient to introduce (as in the exercise on page 36)

$$\not{k} \equiv -i\gamma_\mu k_\mu \quad . \quad (5.59)$$

One can readily verify that

$$(\not{k} - m)(\not{k} + m) = \not{k}^2 - m^2 = -k^2 - m^2$$

and

$$\frac{1}{\not{k} - m} = \frac{\not{k} + m}{(\not{k} - m)(\not{k} + m)} = \frac{-\not{k} - m}{k^2 + m^2} \quad .$$

Thus, (5.58) can also be written as

$$\underbrace{\psi(x)\,\bar{\psi}(0)} = \int \frac{1}{(2\pi)^4} \frac{i}{\not{k} - (m - i\epsilon)} e^{ik\cdot x} d^4x \equiv S_F(x) \quad . \tag{5.60}$$

5.4 Perturbation Series

We begin with (5.20):

$$-\frac{1}{i} \frac{\partial}{\partial t} U(t, t_0) = H_{int}(t)\, U(t, t_0) \quad .$$

Let us replace in the above

$$H_{int} \rightarrow \lambda H_{int} \quad ,$$

and write

$$U(t, t_0) = \sum_{n=0}^{\infty} \lambda^n U_n(t, t_0) \quad . \tag{5.61}$$

Equation (5.20) then becomes

$$-\frac{1}{i} \frac{\partial}{\partial t} \sum_{n=0}^{\infty} \lambda^n U_n(t, t_0) = \lambda H_{int}(t) \sum_{n=0}^{\infty} \lambda^n U_n(t, t_0) \quad . \tag{5.62}$$

By equating the coefficients of λ^n on both sides we have

$$-\frac{1}{i} \frac{\partial}{\partial t} U_0(t, t_0) = 0 \quad , \tag{5.63}$$

and

$$-\frac{1}{i} \frac{\partial}{\partial t} U_n(t, t_0) = H_{int}(t)\, U_{n-1}(t, t_0) \quad \text{for } n \geq 1 \, . \tag{5.64}$$

Similarly, by substituting (5.61) into the initial condition (5.21) we find at $t = t_0$

$$U_0(t_0, t_0) = 1 \tag{5.65}$$

and

$$U_n(t_0, t_0) = 0 \qquad \text{for } n \geqslant 1 \ . \tag{5.66}$$

Equations (5.63) and (5.65) give

$$U_0(t, t_0) = 1 \ . \tag{5.67}$$

The first-order perturbation term, $n = 1$, can be readily obtained by using (5.64), (5.66) and (5.67). The solution is

$$U_1(t, t_0) = -i \int_{t_0}^{t} H_{int}(t') \, dt' \ . \tag{5.68}$$

Likewise, the solution for $n = 2$ is

$$\begin{aligned} U_2(t, t_0) &= (-i)^2 \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \, H_{int}(t_1) \, H_{int}(t_2) \\ &= \frac{(-i)^2}{2!} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \, T(H_{int}(t_1) \, H_{int}(t_2)) \ , \end{aligned} \tag{5.69}$$

etc. By setting $\lambda = 1$ and because of (5.23), we obtain the perturbation series expansion of the S-matrix

$$\begin{aligned} S = U(\infty, -\infty) &= 1 - i \int_{-\infty}^{\infty} H_{int}(t) \, dt \\ &+ \frac{(-i)^2}{2!} \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \, T(H_{int}(t_1) \, H_{int}(t_2)) \\ &+ \frac{(-i)^3}{3!} \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \int_{-\infty}^{\infty} dt_3 \, T(H_{int}(t_1) \, H_{int}(t_2) \, H_{int}(t_3)) \\ &+ \cdots \ , \end{aligned} \tag{5.70}$$

where, as in (5.69), T denotes the time-ordered product. For actual computation, it is useful to convert all these time-ordered products in the series into normal products. The systematics of this conversion is given by Wick's theorem.

5.5 Wick Theorem

We first generalize the definition of the normal product (5.39) to include a c. number multiplicative factor:

$$: c\, X_1 X_2 \cdots X_n : \; = \; c : X_1 X_2 \cdots X_n : \tag{5.71}$$

where c is a c. number. Thus, we have

$$: \underbrace{X_1 X_2} X_3 \cdots X_n : \; = \; \underbrace{X_1 X_2} : X_3 \cdots X_n : \;,$$

$$: \underbrace{X_1 X_2 X_3} \cdots X_n : \; = \; \delta_p \underbrace{X_1 X_3} : X_2 \cdots X_n : \;, \text{ etc.} \tag{5.72}$$

in which the definition of δ_p is the same as that in (5.39); i.e., $\delta_p = -1$ if X_2 and X_3 are fermion operators, otherwise $\delta_p = +1$.

Let the space-time position of the operator $X_i(x_i)$ be, as before, $x_i = (\vec{r}_i\,, it_i)$, and $Y(y)$ be an operator like X_i but at the space-time point $y = (\vec{r}_y\,, it_y)$. We first establish a lemma.

Lemma. If

$$t_y \leqslant t_i \quad \text{for all} \quad i = 1, 2, \cdots, n \;, \tag{5.73}$$

then

$$: X_1 X_2 \cdots X_n : Y \; = \; : X_1 X_2 \cdots \underbrace{X_n Y} :$$

$$+ \; : X_1 X_2 \cdots \underbrace{X_{n-1} X_n Y} : \; + \; \cdots$$

$$+ \; : \underbrace{X_1 X_2 \cdots X_n Y} : \; + \; : X_1 X_2 \cdots X_n Y : \;. \tag{5.74}$$

Proof. (i) In the case that Y is an annihilation operator, we have $: X_i Y : \; = X_i Y$. Since $T(X_i Y) = X_i Y$ on account of (5.73), it follows then that

$$\underbrace{X_i Y} \; = \; T(X_i Y) - : X_i Y : \; = \; 0 \;.$$

Hence, except for the last term $: X_1 X_2 \cdots X_n Y :$, all other terms on

the righthand side of (5.74) are zero. One easily sees that the lemma is correct.

(ii) In the case that Y is a creation operator, we shall prove the lemma in three steps:

(a) Suppose that $X_1, X_2, \cdots, X_n$ are all annihilation operators. We have

$$: X_1 X_2 \cdots X_n : Y = X_1 X_2 \cdots X_n Y \quad , \tag{5.75}$$

and

$$: X_i Y : = \delta_i Y X_i$$

where $\delta_i = -1$ if Y and X_i are fermion operators, otherwise $\delta_i = +1$. Consequently,

$$\begin{aligned} : X_1 X_2 \cdots X_n Y : &= \delta_p : Y X_1 X_2 \cdots X_n : \\ &= \delta_p Y X_1 X_2 \cdots X_n \quad . \end{aligned}$$

In addition, because of the hypothesis (5.73) we also have

$$T(X_i Y) = X_i Y \quad ,$$

and therefore

$$X_i Y = \delta_i Y X_i + \underbrace{X_i Y} \quad ,$$

which leads to

$$X_1 \cdots X_n Y = \delta_n X_1 \cdots X_{n-1} Y X_n + X_1 \cdots X_{n-1} \underbrace{X_n Y} \quad .$$

We can permute Y and X_{n-1} in the first term on the righthand side:

$$\begin{aligned} X_1 \cdots X_n Y = {} & \delta_n \delta_{n-1} X_1 \cdots X_{n-2} Y X_{n-1} X_n \\ & + : X_1 \cdots X_{n-2} \underbrace{X_{n-1} X_n} Y : \\ & + : X_1 \cdots X_{n-1} \underbrace{X_n Y} : \quad . \end{aligned}$$

By repeating this process, we obtain

$$X_1 \cdots X_n Y = \delta_p Y X_1 \cdots X_n + : X_1 \cdots X_{n-1} \underbrace{X_n Y} :$$

$$+ : X_1 \cdots X_{n-2} \underbrace{X_{n-1} X_n} Y : + \cdots + : \underbrace{X_1 \cdots X_n} Y : \ .$$

Since, according to (5.75) the lefthand side of the above expression is the same as that in (5.74), the lemma is established in this case.

(b) Next, we suppose that $X_1, X_2, \cdots, X_j$ are creation operators, but $X_{j+1}, X_{j+2}, \cdots, X_n$ are annihilation operators. Because t_y is $\leqslant t_i$ and Y is a creation operator, we have

$$\underbrace{X_i Y} = 0 \qquad \text{for} \qquad i = 1, 2, \cdots, j$$

and

$$: X_1 X_2 \cdots X_n : Y = X_1 \cdots X_j X_{j+1} \cdots X_n Y$$

$$= X_1 \cdots X_j : X_{j+1} \cdots X_n : Y \ .$$

Since $X_{j+1}, \cdots, X_n$ are all annihilation operators, we can use the result of (a) and apply the lemma to the last term. This gives

$$: X_1 \cdots X_n : Y = X_1 \cdots X_j (: X_{j+1} \cdots X_n Y :$$

$$+ : X_{j+1} \cdots \underbrace{X_n Y} : + : X_{j+1} \cdots \underbrace{X_{n-1} X_n} Y :$$

$$+ \cdots + : \underbrace{X_{j+1} \cdots X_n} Y :) \ .$$

The lemma then follows since $X_1, \cdots, X_j$ are all creation operators.

(c) We now consider the general case that each X_i can be either an annihilation operator or a creation operator. From the definition of the normal product (5.39) we can permute their orders so that on the lefthand side of (5.74) the orders of X_i are arranged to have annihilation operators to the right of creation operators. By using the result of (b) we complete the proof of the lemma.

Wick's theorem deals with the conversion of a time-ordered

product into a sum of normal products.

Theorem.

$$T(X_1 X_2 \cdots X_n) = :X_1 X_2 \cdots X_n: + :\underline{X_1} X_2 \cdots X_n: + :\underline{X_1 X_2} X_3 \cdots X_n: + \cdots + :\underline{X_1} X_2 \underline{X_3} X_4 \cdots X_n: + :\underline{X_1 X_2 X_3} X_4 \cdots X_n: + \cdots . \qquad (5.76)$$

Each term on the righthand side consists of a number of contractions between different pairs of the X_i's, and the righthand side is the sum total of all such terms.

Proof. When $n = 2$, the theorem holds because of the definition of contraction (5.45). Let us assume that for $n \leqslant N$, the theorem is correct. We then consider the time-ordered product when $n = N + 1$.

Among the $N + 1$ operators, $X_1, X_2, \cdots, X_{N+1}$, there must be one X_i whose time t_i is the earliest. Let that X_i be X_{N+1}, i.e. $t_{N+1} \leqslant t_j$ for all $j = 1, 2, \cdots, N$. Hence,

$$T(X_1 X_2 \cdots X_N X_{N+1}) = T(X_1 X_2 \cdots X_N) X_{N+1} .$$

Because of the assumption that the theorem holds for $n = N$, we may convert $T(X_1 X_2 \cdots X_N)$ into a sum of normal products by using (5.76). The result is that $T(X_1 X_2 \cdots X_N) X_{N+1}$ can be written as a sum of terms; each is of the form of the lefthand side of (5.74). By using the lemma, we establish (5.76) for $n = N + 1$. The theorem is then proved by induction.

5.6 Applications

The following example illustrates how Wick's theorem can lead to Feynman diagrams and the evaluation of the S-matrix.

Let the Hamiltonian of a spin-0 field be

$$H = H_0 + H_{int} \quad , \tag{5.77}$$

$$H_0 = \tfrac{1}{2} \int : \Pi^2 + (\vec{\nabla}\phi)^2 + m^2 \phi^2 : d^3r \quad , \tag{5.78}$$

$$H_{int} = \int : -\tfrac{1}{2}\delta m^2 \phi^2 + \frac{1}{3!} g_0 \phi^3 + \frac{1}{4!} f_0 \phi^4 : d^3r \tag{5.79}$$

where, as before, Π is the conjugate momentum, g_0 and f_0 are the unrenormalized coupling constants, m is the physical mass and m_0, defined by the following equation, is the mechanical mass

$$m_0^2 = m^2 - \delta m^2 \quad . \tag{5.80}$$

In the weak coupling, δm^2 is assumed to be of the order of g_0^2 and f_0^2. In the following, we shall stay in the interaction representation and expand $\phi(x)$ in the form of (5.30)-(5.31). Because in (5.78) H_0 is expressed in terms of the normal product, we have

$$H_0 = \sum_{\vec{k}} \omega\, a_{\vec{k}}^{\dagger} a_{\vec{k}} \quad , \tag{5.81}$$

where

$$\omega = \sqrt{\vec{k}^2 + m^2} \quad ,$$

just as in (2.44).

We shall now discuss the scattering process in Fig. 5.2. The initial state is

$$| 1, 2 > \equiv a_1^{\dagger} a_2^{\dagger} | 0 > \quad , \tag{5.82}$$

and the final state is

$$| 1', 2' > \equiv a_1'^{\dagger} a_2'^{\dagger} | 0 > \tag{5.83}$$

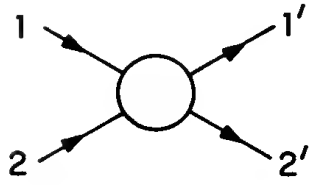


Fig. 5.2. Scattering process $1 + 2 \to 1' + 2'$.

where

$$a_i^\dagger = a^\dagger_{\vec{p}_i} \quad \text{and} \quad a_i'^\dagger = a^\dagger_{\vec{p}_i'} \ ,$$

$i = 1, 2$ and $\vec{p}_i$, $\vec{p}_i'$ are respectively the 3-momenta of particles i and i' . Our purpose is to compute the matrix element $< 1', 2' \mid S \mid 1, 2 >$ where S is the S-matrix given by (5.70).

1. For simplicity, we first discuss the case

$$g_0 = 0 \ , \quad \text{but} \quad f_0 \neq 0 \ . \tag{5.84}$$

To the lowest order of f_0, we have

$$<1', 2' \mid S \mid 1, 2> = \frac{-i}{4!} f_0 \int d^4x <1', 2' \mid :\phi^4(x): \mid 1, 2> . \tag{5.85}$$

This is because δm^2 is $O(f_0^2)$; hence to $O(f_0)$, there is only $\frac{1}{4!} f_0 \phi^4$ in the expression (5.79) for H_{int} . In (5.85) we have to select from $:\phi^4(x):$ only terms proportional to $a_1'^\dagger a_2'^\dagger a_1 a_2$; there are $4! = 24$ such terms due to the different ways of selecting such factors from the four ϕ's . Thus, to $O(f_0)$ we obtain

$$<1', 2' \mid S \mid 1, 2> = -i f_0 \frac{1}{4\sqrt{\omega_1 \omega_2 \omega_1' \omega_2'}} \int \frac{d^4x}{\Omega^2} e^{i(p_1+p_2-p_1'-p_2')\cdot x}$$

where p_1 , p_2 , p_1' and p_2' are respectively the 4-momenta of

particles 1, 2, 1' and 2', the ω_i and ω_i' are the corresponding energies; as before, $x = (\vec{r}, it)$ and $d^4x = d^3r\,dt$. The 4-dimensional integration yields a factor $(2\pi)^4\,\delta^4(p_1 + p_2 - p_1' - p_2')$. It is convenient to introduce $\mathcal{M}$:

$$< 1', 2' \mid S \mid 1, 2 > \equiv \mathcal{M}\frac{(2\pi)^4}{\Omega^2}\,\delta^4(p_1 + p_2 - p_1' - p_2'). \quad (5.86)$$

In this simple example

$$\mathcal{M} = -if_0\,\frac{1}{4\sqrt{\omega_1\omega_2\omega_1'\omega_2'}} \quad . \quad (5.87)$$

In general, $\mathcal{M}$ can be represented by a diagram, or sum of diagrams, called Feynman diagrams. Their rules are called Feynman rules. In Fig. 5.3 we draw the diagram for the $\mathcal{M}$ given by (5.87). The four lines represent the two incoming and two outgoing particles; their intersection is called a 4-point vertex, representing the action of a $:\phi^4:$ interaction.

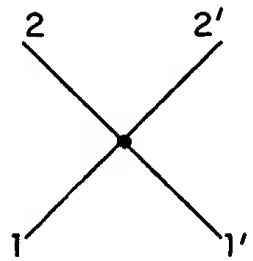


Fig. 5.3. The lowest-order Feynman diagram for $1 + 2 \rightleftarrows 1' + 2'$ when $H_{int} = :\frac{1}{4!}f_0\phi^4:$

The Feynman rule for Fig. 5.3 is: Each external line

gives a factor $\frac{1}{\sqrt{2\omega}}$

where ω is the energy carried by the external line; each 4-point vertex

gives a factor $-i f_0$. (5.88)

The product of all these factors gives the $\mathcal{M}$ of (5.87).

Next, we illustrate the calculation of the cross section. Since

$$\int d^4x\, e^{ipx} = (2\pi)^4 \delta^4(p) \quad ,$$

we have

$$(2\pi)^4 \delta^4(p) \int d^4x\, e^{ipx} = [(2\pi)^4 \delta^4(p)]^2 \quad .$$

On the lefthand side, because of the δ^4-function we may set $p = 0$ in the integrand. The integral becomes

$$\int d^4x = \Omega T$$

where the $\vec{r}$-integration extends over the volume Ω and the time-integration over the interval T. Both Ω and T will $\to \infty$ in the end. Hence we can combine the above two expressions and write

$$[(2\pi)^4 \delta^4(p)]^2 = (2\pi)^4 \delta^4(p)\, \Omega T \tag{5.89}$$

and that leads to, for the square of (5.86),

$$|\langle 1', 2' \,|\, S \,|\, 1, 2\rangle|^2 = |\mathcal{M}|^2 \frac{(2\pi)^4}{\Omega^4} \delta^4(p_1 + p_2 - p_1' - p_2')\, \Omega T \, . \tag{5.90}$$

Let us adopt the laboratory frame in which particle 2 is at rest and particle 1 is moving with velocity $\vec{v}_1$. Assuming that the cross section for $1 + 2 \to 1' + 2'$ is $d\sigma$, we consider a cylinder of volume $v_1 T\, d\sigma$, as shown in Fig. 5.4, where $v_1 = |\vec{v}_1|$. Since our initial

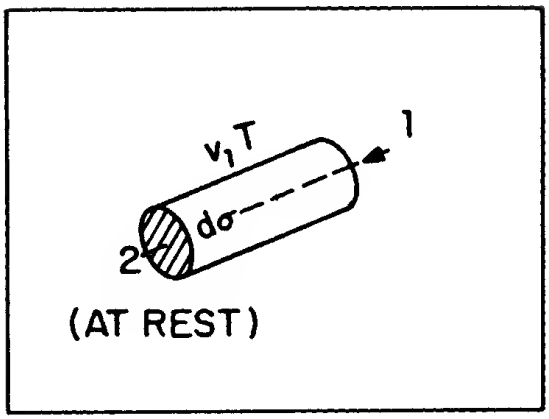


Fig. 5.4. If in the rest frame of particle 2, particle 1 is inside the cylinder $v_1 T d\sigma$, then a reaction $1 + 2 \to 1' + 2'$ will occur in time T.

state (5.82) corresponds to a state in which there are only two particles 1 and 2 in the volume Ω, the probability that particle 1 lies inside this cylinder is

$$\frac{v_1 T d\sigma}{\Omega} \quad . \tag{5.91}$$

If particle 1 is inside this cylinder, then a reaction $1 + 2 \to 1' + 2'$ will occur in time T. Hence, this probability is equal to

$$| < 1', 2' \mid U(\frac{T}{2}, -\frac{T}{2}) \mid 1, 2 > |^2 \tag{5.92}$$

where the U-matrix is given by (5.19). In the limit $\Omega \to \infty$ and then $T \to \infty$, (5.92) becomes (5.90), and therefore upon equating it to (5.91) we have

$$d\sigma = \sum \frac{(2\pi)^4}{\Omega^2 v_1} |\mathcal{M}|^2 \delta^4(p_1 + p_2 - p_1' - p_2') \quad , \tag{5.93}$$

where the sum extends over different 3-momenta $\vec{p}_1'$ and $\vec{p}_2'$ of the

final state. Because of (2.28)

$$\sum_{\vec{p}_i} \to \frac{\Omega}{8\pi^3} \int d^3p_i \ ,$$

we obtain from (5.93)

$$d\sigma = 2\pi \int \frac{d^3p_1' d^3p_2'}{8\pi^3 v_1} \ |\mathcal{m}|^2 \ \delta^4(p_1 + p_2 - p_1' - p_2') \quad (5.94)$$

where v_1 and $\mathcal{m}$ are evaluated in the rest frame of particle 2.

We note that the passage from the definition (5.86) of $\mathcal{m}$ to the above expression for $d\sigma$ has a general validity independent of the special form of the interaction Hamiltonian.

2. We now consider the general case $g_0 \neq 0$ and $f_0 \neq 0$. For simplicity, we shall evaluate the S-matrix element for the reaction $1 + 2 \to 1' + 2'$ only to $O(f_0)$ and $O(g_0^2)$. From (5.70) we find, to these orders and for $|\ 1,2> \neq |\ 1',2'>$,

$$<1',2'|\,S\,|1,2> = -if_0 \frac{1}{4!} \int <1',2'|\,:\phi^4(x):\,|1,2> d^4x$$

$$+ \frac{(-ig_0)^2}{2!\,3!\,3!} \int <1',2'|\,T(\,[:\phi^3(x):][:\phi^3(y):]\,)\,|1,2> d^4x\, d^4y \ . \quad (5.95)$$

We first apply Wick's theorem which converts the T-product in the second integral into a sum of normal products, then retain in this sum only terms proportional to $a_1'^{\dagger} a_2'^{\dagger} a_1 a_2$, as before. There are $(3!)^2$ terms of the form

$$\frac{a_1'^{\dagger} a_2'^{\dagger} a_1 a_2}{4\Omega^2 \sqrt{\omega_1 \omega_2 \omega_1' \omega_2'}} \quad (5.96)$$

multiplied by

$$e^{i(p_1 + p_2)\cdot x - i(p_1' + p_2')\cdot y} + \text{(same terms, but interchanging } x \text{ with } y) \ ; \quad (5.97)$$

there are also $(3!)^2$ terms of the form (5.96) multiplied by

$$e^{i(p_1 - p_1')\cdot x + i(p_2 - p_2')\cdot y} + \text{(same terms, but interchanging } x \text{ with } y) \text{ ;} \tag{5.98}$$

in addition there are $(3!)^2$ terms of the form (5.96) multiplied by

$$e^{i(p_1 - p_2')\cdot x + i(p_2 - p_1')\cdot y} + \text{(same terms, but interchanging } x \text{ with } y) \text{ .} \tag{5.99}$$

In each case the $(3!)^2$ terms and the $x - y$ interchanging terms exactly cancel out the factor $\frac{1}{2!\,3!\,3!}$ outside the integral sign. The three classes (5.97)-(5.99) of terms lead to the three diagrams in Fig. 5.5.

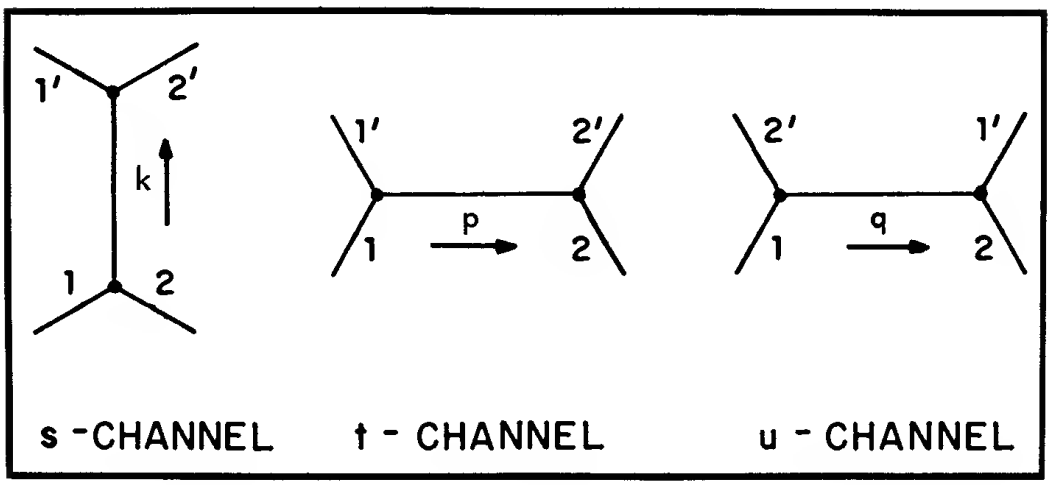


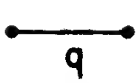
Fig. 5.5. The three $O(g_0^2)$ diagrams for reaction $1 + 2 \to 1' + 2'$. The arrows indicate the directions of momentum flow. *

To evaluate these diagrams we use the following Feynman rules: As in Fig. 5.3 each external line

●——— carries a factor $\frac{1}{\sqrt{2\omega}}$; (5.100)

each internal line

* Because each momentum component can be positive or negative, the arrow directions of these momenta can be arbitrarily drawn.



carries a factor $\dfrac{-i}{q^2 + (m - i\epsilon)^2}$ (5.101)

where q denotes its 4-momentum and $q^2 = \vec{q}^{\,2} - q_0^{\,2}$; each 3-point vertex

gives a factor $-ig_0$. (5.102)

The product of these component-factors is the contribution of the Feynman diagram to $\mathcal{M}$, where $\mathcal{M}$ is defined by (5.86).

By summing over the diagram in Fig. 5.3 and those in Fig. 5.5, we obtain

$$\mathcal{M} = \prod_i \frac{1}{\sqrt{2\omega_i}\sqrt{2\omega_i'}} \left[-if_0 + (-ig_0)^2 \left(\frac{-i}{k^2 + (m - i\epsilon)^2} + \frac{-i}{p^2 + (m - i\epsilon)^2} + \frac{-i}{q^2 + (m - i\epsilon)^2} \right) \right], \tag{5.103}$$

where the subscript i extends over 1 and 2, $k = p_1 + p_2 = p_1' + p_2'$, $p = p_1 - p_1' = p_2' - p_2$ and $q = p_1 - p_2' = p_1' - p_2$.

The differential cross section for the reaction $1 + 2 \to 1' + 2'$ is again given by (5.94).

The Feynman rules (5.88) and (5.100)-(5.102) can be applied to any Feynman diagram of arbitrary order. At each vertex the flow of 4-momenta is conserved. In the high-order diagrams, there will be those that contain loops. An example is given in Fig. 5.6, which is one of the $O(f_0^2)$ diagrams for the scattering process $1 + 2 \to 1' + 2'$.

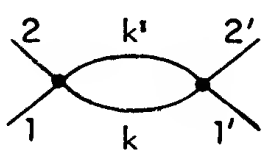


Fig. 5.6. A sample loop diagram.

Neither Fig. 5.3 nor Fig. 5.5 contains any loop; they are called tree diagrams. In a tree diagram, the momenta carried by the internal lines are all determined by the external momenta. However, in a loop diagram each loop carries a free momentum to be integrated over. For example, in Fig. 5.6 the internal 4-momentum k is a free variable, and that gives an additional factor

$$\int \frac{d^4k}{(2\pi)^4} \quad . \tag{5.104}$$

In any loop diagram, there is such a factor for every free momentum.

5.7 Differential Cross Sections for $1 + 2 \to 1' + 2' + \cdots + n'$

1. Equation (5.94) for the differential cross section can be generalized to an n-body final state:

$$1 + 2 \to 1' + 2' + \cdots + n' \quad . \tag{5.105}$$

As in (5.86) we define $\mathcal{M}$ by

$$\langle 1', 2', \cdots n' \mid S \mid 1, 2 \rangle = \frac{(2\pi)^4}{\Omega^{1+(n/2)}} \delta^4\left(\sum_{i=1}^{n} p_i' - p_1 - p_2\right) \mathcal{M} \tag{5.106}$$

where p_i and p_i' denote respectively the 4-momenta of particles i and i' . The process (5.105) can again be expressed as a sum of Feynman diagrams. By using exactly the same Feynman rules (5.88), (5.100)-(5.102) and (5.104), we obtain the contribution of each diagram to $\mathcal{M}$. By following the same considerations given between (5.90) and (5.93), we see that

$$d\sigma = \sum \frac{(2\pi)^4}{\Omega^n v_1} |\mathcal{M}|^2 \delta^4\left(\sum_{i=1}^{n} p_i' - p_1 - p_2\right)$$

where the sum extends over different 3-momenta $\vec{p}_1'$, $\vec{p}_2'$, $\cdots$, $\vec{p}_n'$. When $\Omega \to \infty$, the above expression becomes

$$d\sigma = \frac{2\pi}{v_1}\left(\frac{1}{8\pi^3}\right)^{n-1}\int |\mathcal{M}|^2\,\delta^4\Big(\sum_{i=1}^{n} p_i' - p_1 - p_2\Big)\prod_{i=1}^{n} d^3p_i' \tag{5.107}$$

where all momenta and v_1, the velocity of particle 1, are in the rest frame of 2.

We note that the Feynman rule (5.101) for the internal line is Lorentz-invariant; so are the rules (5.88), (5.102) and (5.104) for the vertices and for the loop-momentum integration. However, the rule (5.100) for the external line is not. Thus, it is convenient to separate out the external-line factors in $\mathcal{M}$ by introducing

$$A \equiv (\sqrt{2})^{n+2}\sqrt{\omega_1'\cdots\omega_n'\omega_1\omega_2}\;\mathcal{M}, \tag{5.108}$$

which is invariant under Lorentz transformation. The differential cross section (5.107) can now be written as

$$d\sigma = \frac{2\pi}{v_1}(\text{phase space})_n \cdot |A|^2 \frac{1}{4\omega_1\omega_2} \tag{5.109}$$

where

$$(\text{phase space})_n \equiv \left(\frac{1}{8\pi^3}\right)^{n-1}\int \prod_{i=1}^{n}\frac{d^3p_i'}{2\omega_i'}\,\delta^4\Big(\sum_{i=1}^{n} p_i' - p_1 - p_2\Big) \tag{5.110}$$

is the Lorentz-invariant n-body phase space.

Sometimes it is convenient to choose the center-of-mass system. We shall show that, in the c.m. system, the differential cross section for reaction (5.105) is given by

$$d\sigma = \frac{2\pi}{|\vec{v}_1 - \vec{v}_2|}\left(\frac{1}{8\pi^3}\right)^{n-1}\int |\mathcal{M}|^2\,\delta^4\Big(\sum_{i=1}^{n} p_i' - p_1 - p_2\Big)\prod_{i=1}^{n} d^3p_i' \tag{5.111}$$

where $\vec{v}_1$ and $\vec{v}_2$ are respectively the velocities of initial particles 1 and 2. [Actually, (5.111) is valid in any frame Σ given in Fig. 5.7.]

Let Σ_{lab} denote the reference frame in which particle 2 is

at rest and particle 1 has velocity $\vec{v}_1$. Consider another frame Σ, as shown in Fig. 5.7, which is moving with a constant velocity $\vec{u} \parallel \vec{v}_1$

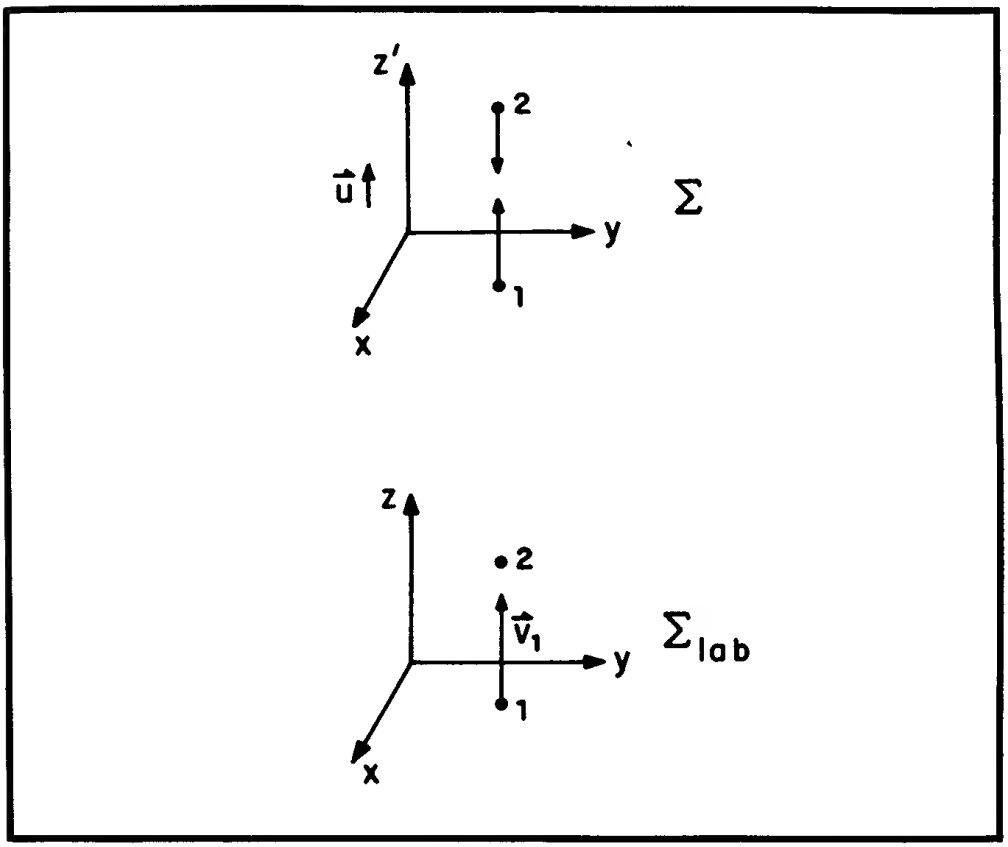


Fig. 5.7. A frame Σ that is moving with a uniform velocity $\vec{u} \parallel \vec{v}_1$ with respect to Σ_{lab}.

with respect to Σ_{lab}. In Σ_{lab}, choose the z-axis to be parallel to $\vec{v}_1$; the 4-momenta of particles 1 and 2 become respectively

$$p_1 = \omega_1(0, 0, v_1, i) \quad \text{and} \quad p_2 = m(0, 0, 0, i) \ .$$

Let $(\omega_i)_\Sigma$ be the energy of particle i in Σ. We have, according to the Lorentz transformation

$$(\omega_1)_\Sigma = \frac{\omega_1 - u v_1 \omega_1}{\sqrt{1 - u^2}} \quad \text{and} \quad (\omega_2)_\Sigma = \frac{m}{\sqrt{1 - u^2}} \ ,$$

which gives

$$(\omega_1 \omega_2)_\Sigma = \omega_1 m \frac{1 - u v_1}{1 - u^2} \quad . \tag{5.112}$$

The velocities of 1 and 2 in Σ are both parallel to the z-axis; their components are respectively

$$(v_1)_\Sigma = \frac{v_1 - u}{1 - u v_1} \qquad \text{and} \qquad (v_2)_\Sigma = -u \quad .$$

Hence we have

$$| \vec{v}_1 - \vec{v}_2 |_\Sigma = (v_1 - v_2)_\Sigma = v_1 \frac{1 - u^2}{1 - u v_1} \quad . \tag{5.113}$$

Throughout, ω_1 and v_1 without the subscript Σ refer to the energy and the velocity of 1 in Σ_{lab}. Combining (5.112) and (5.113), we obtain

$$(\omega_1 \omega_2 | \vec{v}_1 - \vec{v}_2 |)_\Sigma = \omega_1 m v_1 \quad . \tag{5.114}$$

Next, we note that in Σ_{lab}, since $\omega_2 = m$, (5.109) can be written as

$$d\sigma = \frac{2\pi}{v_1} (\text{phase space})_n \, | A |^2 \, \frac{1}{4 \omega_1 m} \quad . \tag{5.115}$$

Since A and $(\text{phase space})_n$ are both Lorentz-invariant, by using (5.114) we derive the expression (5.111) for any such moving frame Σ, and therefore also for $\Sigma_{c.m.}$.

2. The passage from (5.107) for $d\sigma$ in Σ_{lab} to (5.111) in Σ is a general one, valid for particles of arbitrary spin. Therefore it may be useful to give an alternative proof.

In any frame Σ, let ρ_1 and ρ_2 be the number densities of initial particles 1 and 2, Ω the volume of the system and T the total time interval. It is convenient to introduce the reaction rate R, defined by

$$(\rho_1 \rho_2 R)_\Sigma \equiv (\text{total number of reactions}/\Omega T)_\Sigma \quad .$$

Because ΩT is the four-dimensional volume, it is Lorentz invariant, as is the total number of reactions. Consequently, the lefthand side is also Lorentz invariant; i.e.

$$(\rho_1 \rho_2 R)_\Sigma = (\rho_1 \rho_2 R)_{\Sigma'} \tag{5.116}$$

where Σ' denotes any other frame of reference. In any frame Σ, for $\rho_1 = \rho_2 = \frac{1}{\Omega}$, the total number of reactions in ΩT is given by (5.92), except that the bra is now $< 1', 2', \cdots, n' \mid$. Hence in the limit $\Omega \to \infty$ and $T \to \infty$ we have

$$R_\Sigma = \sum \frac{\Omega}{T} \mid < 1', 2', \cdots, n' \mid S \mid 1, 2 > \mid^2$$

where the sum extends over all final 3-momenta. Let $\mathcal{M}$ be related to the matrix element of S by (5.106). The above expression becomes

$$R_\Sigma = \sum \frac{(2\pi)^4}{\Omega^n} \delta^4\Big(\sum_{i=1}^{n} p_i' - p_1 - p_2\Big) \mid \mathcal{M} \mid^2 \quad .$$

Because of (2.28), it can be written as

$$R_\Sigma = \frac{2\pi}{(8\pi^3)^{n-1}} \int \prod_{i=1}^{n} d^3 p_i' \mid \mathcal{M} \mid^2 \delta^4\Big(\sum_{i=1}^{n} p_i' - p_1 - p_2\Big) \tag{5.117}$$

in which all momenta are to be evaluated in the same system Σ. Now, in the laboratory frame we have, on account of (5.91),

$$d\sigma = \left(\frac{R}{v_1}\right)_{\Sigma_{lab}} , \tag{5.118}$$

which gives (5.107).

Let j_1 and j_2 be, respectively, the four-dimensional current vectors of particles 1 and 2. In Σ_{lab}, since particle 2 is at rest, they are given by

$$(j_1)_{\Sigma_{lab}} = \rho_1(\vec{v}_1, i)$$

and

$$(j_2)_{\Sigma_{lab}} = \rho_2(0, i) \quad .$$

Consider a frame Σ which is moving with a constant velocity $\vec{u} \,/\!/\, \vec{v}_1$ with respect to Σ_{lab}. On account of Lorentz transformation, the number densities of particles 1 and 2 in Σ and in Σ_{lab} are related by

$$(\rho_1)_{\Sigma} = \frac{1 - uv_1}{\sqrt{1 - u^2}} (\rho_1)_{\Sigma_{lab}}$$

and

$$(\rho_2)_{\Sigma} = \frac{1}{\sqrt{1 - u^2}} (\rho_2)_{\Sigma_{lab}} \quad .$$

Thus,

$$(\rho_1 \rho_2)_{\Sigma} = \frac{1 - uv_1}{1 - u^2} (\rho_1 \rho_2)_{\Sigma_{lab}} \quad .$$

By using (5.113), we see that

$$(\rho_1 \rho_2 \mid \vec{v}_1 - \vec{v}_2 \mid)_{\Sigma} = (\rho_1 \rho_2 v_1)_{\Sigma_{lab}} \quad ,$$

which together with (5.116) and (5.118) gives

$$d\sigma = \left(\frac{R}{v_1}\right)_{\Sigma_{lab}} = \left(\frac{R}{\mid \vec{v}_1 - \vec{v}_2 \mid}\right)_{\Sigma} \quad . \tag{5.119}$$

Because of (5.117), the expression (5.111) for $d\sigma$ follows.

Problem 5.1. Let the total Hamiltonian of a spin-0 field

$$\phi(\vec{r}, t) = \sum_{\vec{k}} (2\omega\Omega)^{-\frac{1}{2}} (a_{\vec{k}}(t)\, e^{i\vec{k}\cdot\vec{r}} + a^{\dagger}_{\vec{k}}(t)\, e^{-i\vec{k}\cdot\vec{r}})$$

be

$$H = H_0 + \int J(\vec{r})\, \phi(\vec{r}, t)\, d^3r \tag{5.120}$$

where

$$H_0 = \tfrac{1}{2} \int : (\Pi^2 + (\vec{\nabla}\phi)^2 + \mu^2 \phi^2) : d^3r \quad , \tag{5.121}$$

$\omega = (\vec{k}^2 + \mu^2)^{\frac{1}{2}}$, $J(\vec{r}) = J(\vec{r})^*$ is a c. number function independent

of time, and $\Pi(\vec{r}, t)$ is the conjugate momentum of $\phi(\vec{r}, t)$ which satisfies

$$[\Pi(\vec{r}, t), \phi(\vec{r}', t)] = -i\delta^3(\vec{r} - \vec{r}') \quad .$$

(i) Prove that

$$S H S^{\dagger} = H_0 + \Omega \lambda_J$$

where S is a unitary matrix given by

$$S = \exp \sum_{\vec{k}} (-a_{\vec{k}} j_{\vec{k}}^* + a_{\vec{k}}^{\dagger} j_{\vec{k}}) (2\omega^3)^{-\frac{1}{2}} \quad ,$$

$$j_{\vec{k}} = \int \Omega^{-\frac{1}{2}} J(\vec{r}) e^{-i\vec{k}\cdot\vec{r}} d^3r$$

and

$$\lambda_J = -\int (16\pi^3 \omega^2)^{-1} |j_{\vec{k}}|^2 d^3k \quad .$$

[Note: $e^{A+B} = e^A e^B e^{-\frac{1}{2}[A,B]}$ for any two operators A and B whose commutator $[A, B]$ is a c. number.] Thus, the eigenstate $|n_{\vec{k}}>_0 \equiv \prod_{\vec{k}} (n_{\vec{k}}!)^{-\frac{1}{2}} (a_{\vec{k}}^{\dagger})^{n_{\vec{k}}} |\text{vac}>_0$ of H_0, with $n_{\vec{k}} = 0, 1, 2, \cdots$, is related to the corresponding eigenstate $|n_{\vec{k}}>$ of H by $|n_{\vec{k}}> = S^{\dagger} |n_{\vec{k}}>_0$, where

$$H |n_{\vec{k}}> = (\sum_{\vec{k}} n_{\vec{k}} \omega + \Omega \lambda_J) |n_{\vec{k}}> \quad .$$

The new vacuum state $|\text{vac}>$ satisfies $H |\text{vac}> = \Omega \lambda_J |\text{vac}>$.

(ii) Regard $\Omega \lambda_J$ as a functional of $J(\vec{r})$. Show that the variational derivative

$$\frac{\delta \lambda_J}{\delta J(\vec{r})} \equiv \bar{\phi}(\vec{r}) = -\sum_{\vec{k}} \frac{1}{2\omega^2 \Omega^{\frac{1}{2}}} (j_{\vec{k}} e^{i\vec{k}\cdot\vec{r}} + j_{\vec{k}}^* e^{-i\vec{k}\cdot\vec{r}}) \quad .$$

Furthermore, $\bar{\phi}(\vec{r})$ is the vacuum expectation of $\phi(\vec{r}, t)$:

$$\bar{\phi}(\vec{r}) = <0 | \phi(\vec{r}, t) | 0> \quad .$$

Note that because $J(\vec{r})$ is assumed to be independent of t, so is $\bar{\phi}(\vec{r})$.

(iii) Show that

$$\mathcal{E}(\bar{\phi}) \equiv \Omega\lambda_J - \int J(\vec{r})\,\bar{\phi}(\vec{r})\,d^3r$$

is the minimum of $<|H_0|>$, taken among all states $|>$ under the constraint

$$<|\phi(\vec{r}, t)|> = \bar{\phi}(\vec{r}) \quad .$$

Problem 5.2. Replace (5.120) in Problem 5.1 by

$$H = H_0 + H_{int}$$

where H_0 remains given by (5.121), but

$$H_{int} = \int (J\phi + \tfrac{1}{2}m^2\phi^2)\,d^3r$$

where, as before, $J(\vec{r}) = J(\vec{r})^*$ is a c. number function. Regard H_{int} as the interaction Hamiltonian. In the interaction representation the propagator $\overset{p}{\text{———}}$ is $-i(\vec{p}^2 - p_0^2 + \mu^2 - i\epsilon)^{-1}$ where $\epsilon = 0+$, and there are vertices ———•——— $= -im^2$, •———→ $\overset{k}{} = -ij_k$ and ———→• $\overset{k}{} = -ij_k^*$, where the arrow indicates the flow direction of the momentum k, whose fourth component is 0 because the c. number function $J(\vec{r})$ is assumed to be independent of time. By summing over graphs, show that the full propagator is

$$D_F(p) \equiv \overset{p}{\text{———}} + \overset{p}{\text{———•———}}\overset{p}{} + \overset{p}{\text{———•———•———}}\overset{p}{} + \cdots$$

$$= -i(\vec{p}^2 - p_0^2 + \mu^2 + m^2 - i\epsilon)^{-1}$$

and the new vacuum energy is $\Omega(\Lambda_J + \Delta)$ where

$$-i\Lambda_J = \text{•———•} + \text{•———•———•} + \text{•———•———•———•} + \cdots$$

$$= i\tfrac{1}{2}\int [8\pi^3(\vec{k}^2 + \mu^2 + m^2)]^{-1}\,|j_{\vec{k}}|^2\,d^3k$$

and

$$-i\Delta = \quad + \quad + \quad + \cdots$$

$$= -i\tfrac{1}{2}\int (8\pi^3)^{-1}\,[(\vec{k}^2+\mu^2+m^2)^{\frac{1}{2}} - (\vec{k}^2+\mu^2)^{\frac{1}{2}}]\,d^3k\,.$$

Notice that the graphs , , , etc. all have a symmetry number* factor $\frac{1}{2}$.

Problem 5.3. (i) Consider the following matrix element of the operator

$$J_\mu \equiv :i\psi_b^\dagger \gamma_4 \gamma_\mu (C_V + C_A \gamma_5)\psi_a : \tag{5.122}$$

$$<J_\mu> \equiv <b \mid J_\mu \mid a> \tag{5.123}$$

where ψ_a and ψ_b are both spin-$\frac{1}{2}$ quantized fields, C_V and C_A are constants, $\mid a>$ is the free a-particle state of helicity s_a, mass m_a and 4-momentum a_μ, and $\mid b>$ the free b-particle state of helicity s_b, mass m_b and 4-momentum b_μ.

$$T_{\mu\nu} \equiv \pm \sum_{s_a, s_b} <J_\mu><J_\nu>^* \tag{5.124}$$

where + is for $\nu \neq 4$ and - is for $\nu = 4$. Show that by using the exercise on page 36 and setting the volume $\Omega = 1$,

$$a_0 b_0 T_{\mu\nu} = (|C_V|^2 + |C_A|^2)(a_\mu b_\nu + a_\nu b_\mu - \delta_{\mu\nu}\, a\cdot b)$$
$$+ (C_V^* C_A + C_A^* C_V)\epsilon_{\mu\nu\lambda\delta}\, a_\lambda b_\delta - m_a m_b (|C_V|^2 - |C_A|^2)\delta_{\mu\nu} \tag{5.125}$$

where $a_0 = (\vec{a}^2 + m_a^2)^{\frac{1}{2}}$ and $b_0 = (\vec{b}^2 + m_b^2)^{\frac{1}{2}}$, $\delta_{\mu\nu} = 1$ if $\mu = \nu$, but $= 0$ if $\mu \neq \nu$, while $\epsilon_{\mu\nu\lambda\delta} = 1$ or -1 depending on whether $\mu\nu\lambda\delta$ is an even or odd permutation of 1234, and $= 0$ otherwise.

---

* See page 500 for the definition of symmetry number. Note that each Hermitian boson line is unarrowed. [In contrast, the line of a charged (complex) field carries an arrow pointing in the direction of the charge flow. Cf. also pages 530-32 and 90.]

(ii) Show that (5.125) remains valid, if instead of (5.123)

$$< J_\mu > = < \bar{a} \mid J_\mu \mid \bar{b} >$$

where $\mid \bar{a} >$ is the free $\bar{a}$-particle state of 4-momentum a_μ and $\mid \bar{b} >$ the free $\bar{b}$-particle state of 4-momentum b_μ.

(iii) If, instead of Eq. (5.123)

$$< J_\mu > = < 0 \mid J_\mu \mid a\bar{b} > \qquad \text{or} \qquad < \bar{a}b \mid J_\mu \mid 0 > \ ,$$

then $T_{\mu\nu}$, defined by (5.124), remains given by (5.125), except for the change

$$- m_a m_b \rightarrow + m_a m_b$$

on the righthand side, where $\mid a\bar{b} >$, or $\mid \bar{a}b >$, is the state of a free a-particle (or $\bar{a}$-particle) of 4-momentum a_μ together with a free $\bar{b}$-particle (or b-particle) of 4-momentum b_μ.

Problem 5.4. Phenomenologically, the weak interaction Lagrangian for

$$\nu_\ell + a \rightarrow \ell^- + b \tag{5.126}$$

and

$$\bar{\nu}_\ell + b \rightarrow \ell^+ + a \tag{5.127}$$

can be written as

$$\mathcal{L} = 2^{-\frac{1}{2}} [j_\mu J_\mu + j_\mu^\dagger J_\mu^\dagger]$$

where a and b are some spin-$\frac{1}{2}$ hadrons, $\ell^\pm$ denotes the charged lepton $e^\pm$ or $\mu^\pm$,

$$j_\mu = i \psi_\ell^\dagger \gamma_4 \gamma_\mu (1 + \gamma_5) \psi_{\nu_\ell}$$

and

$$J_\mu = i \psi_b^\dagger \gamma_4 \gamma_\mu (C_V + C_A \gamma_5) \psi_a \ .$$

Neglect the strong interaction of a and b, assume the hadron mass $m_a = m_b = m$, and set the lepton mass $m_\ell = 0$. Show that to the

lowest order in C_V and C_A, the differential cross sections for reactions (5.126) and (5.127) are respectively

$$\frac{d\sigma_\nu}{dy} = \frac{mE_\nu}{2\pi}\left[\,|C_V + C_A|^2 + |C_V - C_A|^2 (1-y)^2 + (|C_A|^2 - |C_V|^2)\,\frac{m}{E_\nu}\,y\,\right]$$

and (5.128)

$$\frac{d\sigma_{\bar\nu}}{dy} = \frac{mE_\nu}{2\pi}\left[\,|C_V - C_A|^2 + |C_V + C_A|^2 (1-y)^2 + (|C_A|^2 - |C_V|^2)\,\frac{m}{E_\nu}\,y\,\right]$$

where

$$y \equiv \frac{E_\nu - E_\ell}{E_\nu}\ ,$$

E_ν and E_ℓ are, respectively, the energies of the neutrino and $\ell^\pm$ in the laboratory frame (i.e., the rest frame of the initial hadron).

Note that in any frame $y = q\cdot a/k\cdot a$ for (5.126) and $q\cdot b/k\cdot b$ for (5.127) where $q_\mu \equiv k_\mu - k'_\mu$ and k_μ, k'_μ, a_μ, b_μ are, respectively, the 4-momenta of the neutrino, $\ell^\pm$, a and b, the range of y is from 0 to 1. For further discussions see Chapters 21 and 23.

References

F. Dyson, Phys.Rev. 75, 486 (1949).

R. P. Feynman and A. R. Hibbs, Quantum Mechanics and Path Integrals (New York, McGraw-Hill, 1965).

G. C. Wick, Phys.Rev. 80, 268 (1950).

See also Chapter 19 for a discussion of the path-integration method, which is the original way Feynman invented his diagrams.

Chapter 6

QUANTUM ELECTRODYNAMICS

In quantum electrodynamics we consider the electromagnetic interaction between photons, electrons and positrons. If we wish, we may also include other charged leptons, such as $\mu^{\pm}$ and $\tau^{\pm}$. Because the photon is of spin 1 and mass 0, this also serves as an example of how to deal with the $m = 0$ limit of a vector field.

6.1 Lagrangian

Let ψ be the electron field, A_μ be the electromagnetic 4-potential and $F_{\mu\nu}$ be the electromagnetic field tensor, which is related to A_μ by

$$F_{\mu\nu} \equiv \frac{\partial}{\partial x_\mu} A_\nu - \frac{\partial}{\partial x_\nu} A_\mu \quad . \tag{6.1}$$

The Lagrangian density in quantum electrodynamics is

$$\mathcal{L} = \mathcal{L}_e + \mathcal{L}_\gamma + \mathcal{L}_{int} \quad , \tag{6.2}$$

where

$$\mathcal{L}_\gamma = -\tfrac{1}{4} F_{\mu\nu}^2 = \tfrac{1}{2}(\vec{E}^2 - \vec{B}^2) \quad , \tag{6.3}$$

$$\mathcal{L}_e = -\psi^\dagger \gamma_4 (\gamma_\mu \frac{\partial}{\partial x_\mu} + m)\, \psi \tag{6.4}$$

and

$$\mathcal{L}_{int} = e\, j_\mu A_\mu + \psi^\dagger \gamma_4 \psi\, \delta m \quad . \tag{6.5}$$

In (6.3), $\vec{E}$ is the electric field and $\vec{B}$ the magnetic field, related

to $A_\mu = (\vec{A}, iA_0)$ by

$$\vec{E} \equiv -\vec{\nabla} A_0 - \dot{\vec{A}} \ , \qquad \vec{B} \equiv \vec{\nabla} \times \vec{A} \ . \tag{6.6}$$

In (6.5), the electromagnetic current is

$$j_\mu \equiv i \psi^\dagger \gamma_4 \gamma_\mu \psi \ , \tag{6.7}$$

e is the unrenormalized charge and δm is defined to be the difference between the physical mass m and the mechanical mass m_0 :

$$\delta m = m - m_0 \ . \tag{6.8}$$

From the variational principle (2.10), we obtain the equations of motion for A_μ and ψ :

$$\frac{\partial}{\partial x_\mu} F_{\mu\nu} = -e j_\nu \tag{6.9}$$

and

$$\gamma_\mu \left(\frac{\partial}{\partial x_\mu} - ie A_\mu \right) \psi + m_0 \psi = 0 \ . \tag{6.10}$$

6.2 Coulomb Gauge

The Lagrangian density is invariant under the gauge transformation

$$A_\mu \rightarrow A_\mu + \frac{\partial \theta}{\partial x_\mu} \tag{6.11}$$

and

$$\psi \rightarrow e^{ie\theta} \psi \ . \tag{6.12}$$

By choosing a suitable function θ, we may impose the transversality condition on $\vec{A}$:

$$\vec{A} = \vec{A}^{tr} \ , \qquad \text{i.e.,} \quad \vec{\nabla} \cdot \vec{A} = 0 \ . \tag{6.13}$$

This particular choice is called the Coulomb gauge. Since the theory is gauge invariant, any choice of gauge should lead to the same physical results. The Coulomb gauge is, however, particularly convenient for the purpose of quantization, as we shall see. Since the Lagrangian

density (6.2) does not contain $\dot{A}_0$, we shall regard A_0 as a dependent variable, just as in Chapter 4. When $\nu = 4$, (6.9) becomes simply

$$\vec{\nabla} \cdot \vec{E} = e\rho \tag{6.14}$$

where $j_4 = i\rho = i\psi^\dagger\psi$. Because of (6.6) and (6.13), the above equation can be written as

$$\nabla^2 A_0 = -e\rho \ . \tag{6.15}$$

In the Coulomb gauge, we regard A_0 as a functional of $\psi^\dagger\psi$, given by the solution of the Laplace equation

$$A_0(\vec{r}, t) \equiv \int \frac{e\rho(\vec{r}', t)}{4\pi\,|\vec{r} - \vec{r}'|}\, d^3r' \ . \tag{6.16}$$

We may decompose the electric field in (6.6) into two terms:

$$\vec{E} = \vec{E}^{tr} + \vec{E}^{\ell} \tag{6.17}$$

in which the transverse component is

$$\vec{E}^{tr} = -\dot{\vec{A}} \tag{6.18}$$

and the longitudinal component is

$$\vec{E}^{\ell} = -\vec{\nabla} A_0 \ . \tag{6.19}$$

Clearly, $\vec{E}^{\ell}$ is irrotational and $\vec{E}^{tr}$, because of (6.13), is divergence free. Through partial integration, the following volume integral can be reduced to 0 :

$$\int \vec{E}^{tr} \cdot \vec{E}^{\ell}\, d^3r = \int \dot{\vec{A}} \cdot \vec{\nabla} A_0\, d^3r = 0 \ .$$

Therefore we may replace the Lagrangian density $\mathcal{L}_\gamma$ in (6.3) by

$$\mathcal{L}_\gamma = \tfrac{1}{2}\left[(\vec{E}^{tr})^2 + (\vec{E}^{\ell})^2 - \vec{B}^2 \right] \ , \tag{6.20}$$

without changing the resulting Lagrangian $L = \int \mathcal{L}\, d^3r$. The

conjugate momentum $\vec{\Pi}$ of the electromagnetic potential $\vec{A}$ is

$$\vec{\Pi} \equiv \frac{\partial \mathcal{L}}{\partial \dot{\vec{A}}} = -\vec{E}^{tr} \quad , \tag{6.21}$$

and the conjugate momentum of the electron field is

$$\mathcal{P} \equiv \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = i\psi^{\dagger} \quad . \tag{6.22}$$

By following the usual canonical procedure, we find the Hamiltonian density to be

$$\mathcal{H} = \mathcal{H}_{\gamma} + \mathcal{H}_{e} + \mathcal{H}_{int} \quad , \tag{6.23}$$

where

$$\mathcal{H}_{\gamma} = \tfrac{1}{2}(\vec{E}^{tr})^2 + \tfrac{1}{2}\vec{B}^2 \quad , \tag{6.24}$$

$$\mathcal{H}_{e} = \psi^{\dagger}(-i\vec{\alpha}\cdot\vec{\nabla} + \beta m)\psi \tag{6.25}$$

and

$$\mathcal{H}_{int} = -\psi^{\dagger}\beta\psi \cdot \delta m - \tfrac{1}{2}(\vec{E}^{\ell})^2 - e j_{\mu} A_{\mu} \quad . \tag{6.26}$$

It is convenient to separate out from $\mathcal{H}_{int}$ a part that corresponds to the Coulomb interaction between the charge density. We define

$$\mathcal{H}_{Coul} \equiv -\tfrac{1}{2}(E^{\ell})^2 + e\rho A_0 \quad .$$

Because of (6.15)-(6.16) and (6.19), the space integral of $\mathcal{H}_{Coul}$ is

$$\begin{aligned} H_{Coul} &= \int \mathcal{H}_{Coul}\, d^3r = \int (\tfrac{1}{2} A_0 \nabla^2 A_0 + e\rho A_0)\, d^3r \\ &= \tfrac{1}{2}\int e\rho A_0\, d^3r \\ &= \tfrac{1}{2}\int \frac{e^2 \rho(\vec{r},t)\,\rho(\vec{r}',t)}{4\pi\,|\vec{r}-\vec{r}'|}\, d^3r\, d^3r' \end{aligned} \tag{6.27}$$

The interaction Hamiltonian is then given by

$$\begin{aligned} H_{int} &= \int \mathcal{H}_{int}\, d^3r \\ &= -\int \psi^{\dagger}\beta\psi \cdot \delta m\, d^3r + H_{Coul} - e\int \vec{j}\cdot\vec{A}\, d^3r \end{aligned} \quad . \tag{6.28}$$

The total Hamiltonian H is given by the space integral of (6.23). In H, the generalized coordinates are $\vec{A} = \vec{A}^{tr}$ and ψ; the generalized momenta are $\vec{\Pi} = -\vec{E}^{tr}$ and $\mathcal{P} = i\psi^\dagger$.

6.3 Quantization

The quantization procedure in the Coulomb gauge can be carried out in a straightforward manner. Because of the transversality condition (6.13), the equal-time commutator between $\vec{\Pi}$ and $\vec{A}$ is

$$[\Pi_i(\vec{r}, t), A_j(\vec{r}', t)] = -i(\delta_{ij} - \nabla^{-2}\nabla_i \nabla_j)\,\delta^3(\vec{r} - \vec{r}') \tag{6.29}$$

where the factor $(\delta_{ij} - \nabla^{-2}\nabla_i \nabla_j)$ is to insure that the righthand side satisfies the same divergence-free constraints,

$$\vec{\nabla} \cdot \vec{A} = 0 \quad \text{and} \quad \vec{\nabla} \cdot \vec{\Pi} = 0 \ . \tag{6.30}$$

Between the electron field ψ and its Hermitian conjugate $\psi^\dagger$ we have the usual equal-time anticommutator

$$\{\psi(\vec{r}, t), \psi^\dagger(\vec{r}', t)\} = \delta^3(\vec{r} - \vec{r}') \ . \tag{6.31}$$

Likewise, the equal-time anticommutator between ψ (and that between $\psi^\dagger$) at different space-positions is 0, and so are the equal-time commutators between other pairs of field operators.

At any given time t we may expand $\vec{A}(\vec{r}, t)$ and $\vec{\Pi}(\vec{r}, t)$ in terms of the Fourier series

$$\vec{A}(\vec{r}, t) = \sum_{\vec{k}} \frac{1}{\sqrt{2\omega\Omega}} [\vec{a}_{\vec{k}}(t)\, e^{i\vec{k}\cdot\vec{r}} + \text{h.c.}] \tag{6.32}$$

and

$$\vec{\Pi}(\vec{r}, t) = -\vec{E} = \sum_{\vec{k}} \sqrt{\frac{\omega}{2\Omega}} [-i\vec{a}_{\vec{k}}(t)\, e^{i\vec{k}\cdot\vec{r}} + \text{h. c.}] \tag{6.33}$$

where $\omega = |\vec{k}|$. Because of (6.30), the $\vec{a}_{\vec{k}}$'s satisfy

$$\vec{a}_{\vec{k}}(t) \cdot \vec{k} = 0 \ . \tag{6.34}$$

The expansions (6.32) and (6.33) are valid in any representation. In the interaction representation, the $\vec{a}_{\vec{k}}$'s have the same time dependence as that in the free field. We have

$$\vec{a}_{\vec{k}}(t) \propto e^{-i\omega t} . \tag{6.35}$$

As in Chapter 4, we may introduce for any given $\vec{k}$ a set of three orthogonal unit vectors $\hat{k} = \vec{k}/|\vec{k}|$, $\hat{e}_1$ and $\hat{e}_2$:

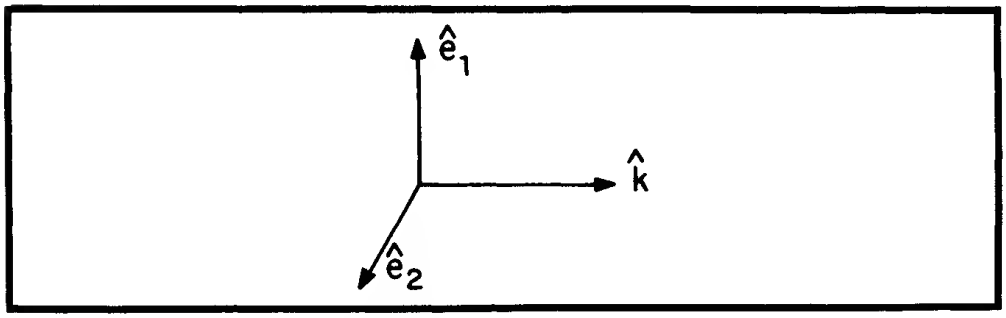


Fig. 6.1. A righthanded orthonormal set of three vectors.

It is convenient to define

$$a^{\dagger}_{\vec{k},s=\pm 1} \equiv \frac{1}{\sqrt{2}} \vec{a}^{\dagger}_{k} \cdot (\hat{e}_1 \pm i\hat{e}_2) , \tag{6.36}$$

and its Hermitian conjugate

$$a_{\vec{k},s=\pm 1} \equiv \frac{1}{\sqrt{2}} \vec{a}_{k} \cdot (\hat{e}_1 \mp i\hat{e}_2) . \tag{6.37}$$

From (6.29) and

$$[A_i(\vec{r}, t), A_j(\vec{r}', t)] = [\Pi_i(\vec{r}, t), \Pi_j(\vec{r}', t)] = 0 ,$$

one can readily verify that

$$[a_{\vec{k},s}(t), a^{\dagger}_{\vec{k}',s'}(t)] = \delta_{\vec{k},\vec{k}'} \delta_{ss'} , \tag{6.38}$$

and

$$[a_{\vec{k},s}(t), a_{\vec{k}',s'}(t)] = 0 \tag{6.39}$$

where s and s' can be $+1$ or -1. It can also be shown that the subscript s denotes the helicity of the photon (= its spin component along the direction of motion). Thus $a^{\dagger}_{\vec{k},s}$ is the creation operator of a photon with momentum $\vec{k}$ and helicity s; its Hermitian conjugate $a_{\vec{k},s}$ is the corresponding annihilation operator.

The S-matrix can be derived by following the steps discussed in Chapter 5. Its perturbation series is given by (5.70) in which the interaction Hamiltonian is the integral of the normal product of $\mathcal{H}_{int}$,

$$H_{int} = \int : \mathcal{H}_{int} : d^3r \tag{6.40}$$

where $\mathcal{H}_{int}$ is given by (6.26). Likewise, the unperturbed Hamiltonian is

$$H_0 = \int : \mathcal{H}_{\gamma} + \mathcal{H}_e : d^3r \tag{6.41}$$

where $\mathcal{H}_{\gamma}$ and $\mathcal{H}_e$ are given by (6.24) and (6.25).

6.4 Photon Propagator and Relativistic Invariance

In the Coulomb gauge, while one can easily carry out the quantization procedure, the Lorentz-invariant character of QED is less obvious, but will be demonstrated using Feynman diagrams.

Let us consider the diagram shown in Fig. 6.2. The amplitude

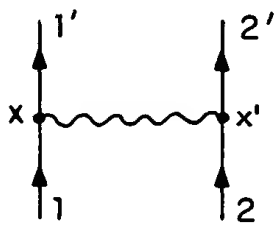


Fig. 6.2. A diagram for $1 + 2 \rightarrow 1' + 2'$.

of this diagram consists of two terms. The first is due to

$$-i\int H_{Coul}\,dt = -\frac{i}{2}\int \frac{e^2}{4\pi}\,\rho(\vec{r},t)\,\rho(\vec{r}',t)\,\frac{d^3r\,d^3\vec{r}'}{|\vec{r}-\vec{r}'|}\,dt$$

$$= \frac{i}{2}\int \frac{e^2}{4\pi}\,j_4(x)\,j_4(x')\,\frac{\delta(t-t')}{|\vec{r}-\vec{r}'|}\,d^4x\,d^4x'\,, \tag{6.42}$$

where $j_4 = i\rho$. The second term is due to

$$\frac{(-i)^2}{2!}\int T(e^2\,\vec{j}(x)\cdot\underbrace{\vec{A}(x)\,\vec{j}(x')\cdot\vec{A}(x')}\,)\,d^4x\,d^4x' \tag{6.43}$$

where T denotes the time-ordered product. Their sum is

$$\frac{(-i)^2}{2!}\int e^2\,T(j_\mu(x)\,D^{Coul}_{\mu\nu}(x-x')\,j_\nu(x'))\,d^4x\,d^4x'\,, \tag{6.44}$$

where $D^{Coul}_{\mu\nu}(x-x')$ is the photon propagator in the Coulomb gauge:

$$x \sim\sim x' = D^{Coul}_{\mu\nu}(x-x') = \begin{cases} D^{tr}_{ij}(x-x') = \underbrace{A_i(x)\,A_j(x')} & \\ \qquad\qquad \text{if} \quad \mu = i \neq 4 & \\ \qquad\qquad \text{and} \quad \nu = j \neq 4\,, & \\ \dfrac{-i\,\delta(t-t')}{4\pi\,|\vec{r}-\vec{r}'|} \quad \text{if} \quad \mu=\nu=4\,, & \\ 0 \quad \text{otherwise}\,. & \end{cases} \tag{6.45}$$

Because $A_i = A_i^{tr}$, in the above expression when $\mu = i \neq 4$ and $\nu = j \neq 4$ the photon propagator is given by

$$D^{tr}_{ij}(x) = \frac{-i}{(2\pi)^4}\int \frac{d^4k}{k^2 - i\epsilon}\left(\delta_{ij} - \frac{k_i\,k_j}{\vec{k}^2}\right) e^{ik\cdot x}\,, \tag{6.46}$$

in which the factor $\left(\delta_{ij} - \dfrac{k_i\,k_j}{\vec{k}^2}\right)$ has the same origin as the factor $(\delta_{ij} - \nabla^{-2}\,\nabla_i\,\nabla_j)$ in (6.29), so that $\nabla_i\,D^{tr}_{ij}(x) = 0$. As before,

$$\epsilon = 0+ \qquad \text{and} \qquad k^2 = \vec{k}^2 - k_0^2\,.$$

We note that

$$\frac{1}{8\pi^3}\int \frac{1}{\vec{k}^2}\, e^{i\vec{k}\cdot\vec{r}}\, d^3k = \frac{2\pi}{8\pi^3}\int e^{ikr\cos\theta}\, dk\, d\cos\theta$$

$$= \frac{1}{4\pi^2}\int_0^\infty \frac{e^{ikr} - e^{-ikr}}{ikr}\, dk = \frac{1}{4\pi r}\,.$$

Thus, in the momentum space the photon propagator becomes

$$D_{\mu\nu}^{\text{Coul}}(k) = \begin{cases} \dfrac{-i}{k^2 - i\epsilon}\left(\delta_{ij} - \dfrac{k_i k_j}{\vec{k}^2}\right) & \text{if } \mu = i \neq 4 \\ & \text{and } \nu = j \neq 4\,, \\ -\dfrac{i}{\vec{k}^2} & \text{if } \mu = \nu = 4\,, \\ 0 \quad \text{otherwise} & \end{cases} \tag{6.47}$$

where

$$D_{\mu\nu}^{\text{Coul}}(k) \equiv \int D_{\mu\nu}^{\text{Coul}}(x)\, e^{-ik\cdot x}\, d^4x\,.$$

According to (6.44) the amplitude for the Feynman diagram in Fig. 6.2 is given by

$$\frac{1}{2!}(-ie)^2 \int d^4x\, d^4x' \left[\langle 1' \mid j_\mu(x) \mid 1\rangle\, D_{\mu\nu}^{\text{Coul}}(x-x')\, \langle 2' \mid j_\mu(x') \mid 2\rangle\right.$$

$$\left. + \text{same terms, but interchanging } x \text{ with } x'\right]\,. \tag{6.48}$$

Let p_i and p_i' be respectively the 4-momenta of particles i and i'. The matrix element $\langle 1' \mid j_\mu(x) \mid 1\rangle$ is proportional to $e^{i(p_1 - p_1')\cdot x}$; likewise $\langle 2' \mid j_\mu(x') \mid 2\rangle$ is proportional to $e^{i(p_2 - p_2')\cdot x'}$. Thus, we may write

$$\langle 1' \mid j_\mu(x) \mid 1\rangle \equiv e^{i(p_1 - p_1')\cdot x}\, a_\mu\,,$$

and

$$\langle 2' \mid j_\mu(x') \mid 2\rangle \equiv e^{i(p_2 - p_2')\cdot x'}\, b_\mu \tag{6.49}$$

where a_μ and b_μ are independent of x and x'. Due to momentum conservation, we have

$$k \equiv p_1 - p_1' = -p_2 + p_2'\,. \tag{6.50}$$

Because of current conservation, we also have

$$\frac{\partial j_\mu(x)}{\partial x_\mu} = 0 \ ,$$

which leads to, on account of (6.49)-(6.50),

$$k_\mu a_\mu = k_\nu b_\nu = 0 \ . \tag{6.51}$$

Transforming to momentum space, we can re-label Fig. 6.2 in the following form:

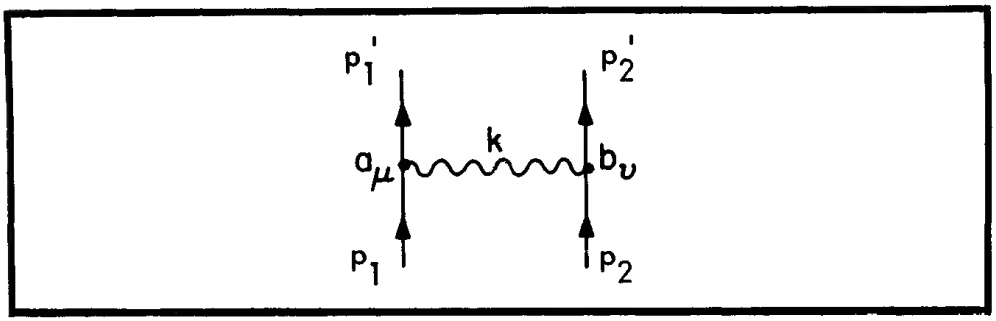


Fig. 6.3. Diagram for $1 + 2 \rightarrow 1' + 2'$ in momentum space.

Because of (6.47), its amplitude is proportional to

$$a_\mu D^{\text{Coul}}_{\mu\nu}(k) b_\nu = -i \left[\frac{1}{k^2} (\vec{a}\cdot\vec{b} - \frac{(\vec{k}\cdot\vec{a})(\vec{k}\cdot\vec{b})}{k^2}) - \frac{a_0 b_0}{k^2} \right] \tag{6.52}$$

where $a_\mu = (\vec{a}, i a_0)$, $b_\mu = (\vec{b}, i b_0)$ and, as before, $k^2 = \vec{k}^2 - k_0^2$. For simplicity we have omitted $-i\epsilon$ in the denominator of $\frac{1}{k^2}$. Since $\vec{k}\cdot\vec{a} = k_0 a_0$ and $\vec{k}\cdot\vec{b} = k_0 b_0$, in accordance with (6.51), the amplitude (6.52) becomes

$$-i \left[\frac{\vec{a}\cdot\vec{b}}{k^2} - \frac{a_0 b_0}{k^2} \left(\frac{k_0^2}{k^2} + 1 \right) \right] = \frac{-i}{k^2} (\vec{a}\cdot\vec{b} - a_0 b_0) \ .$$

Consequently, we can replace the non-covariant Coulomb propagator

$D_{\mu\nu}^{Coul}$ by the covariant propagator $D_{\mu\nu}$:

$$a_\mu D_{\mu\nu}^{Coul}(k)\, b_\nu = a_\mu D_{\mu\nu}(k)\, b_\nu \tag{6.53}$$

where

$$D_{\mu\nu}(k) = \frac{-i}{k^2 - i\epsilon}\left[\delta_{\mu\nu} + \lambda \frac{k_\mu k_\nu}{k^2}\right] . \tag{6.54}$$

Because of (6.51), the Feynman amplitude (6.53) is independent of the parameter λ. The choice of λ is therefore arbitrary. When $\lambda = 0$, it is referred to as the Feynman gauge; when $\lambda = -1$, as the Landau gauge.

It is possible to prove that this replacement is valid in all Feynman diagrams, and thereby establish the Lorentz invariance of the theory.

6.5 Remarks

Because positronium states are all unstable, quantum electrodynamics is one example in which no stable bound state exists. Therefore, in the notation of (6.23)-(6.26) the spectra of the total Hamiltonian

$$H = \int : \mathcal{H}_\gamma + \mathcal{H}_e + \mathcal{H}_{int} : d^3r$$

and (6.55)

$$H_0 = \int : \mathcal{H}_\gamma + \mathcal{H}_e : d^3r$$

are the same, provided the mass m in H_0 is the physical mass of the electron. Any eigenstate $|\, n >$ of the free Hamiltonian,

$$H_0 \,|\, n > = E_n \,|\, n > \ ,$$

can be written as

$$|\, n > = \Pi\, a^\dagger_{\vec{p}_i, s_i}\, b^\dagger_{\vec{p}_j, s_j}\, a^\dagger_{\vec{p}_k, s_k} \,|\, 0 > \tag{6.56}$$

where $a^{\dagger}_{\vec{p}_i, s_i}$, $b^{\dagger}_{\vec{p}_j, s_j}$ and $a^{\dagger}_{\vec{p}_k, s_k}$ are respectively the creation operators of the electron, positron and photon. The state $|0\rangle$ is the vacuum state of H_0,

$$H_0 |0\rangle = 0 . \tag{6.57}$$

Let $U(t, t_0)$ be the solution of (5.20)-(5.21). It is useful to introduce

$$|n^{in}\rangle \equiv U(t, -\infty) |n\rangle \tag{6.58}$$

and

$$|n^{f}\rangle \equiv U^{\dagger}(\infty, t) |n\rangle , \tag{6.59}$$

where t can be any finite time, the superscripts in and f denote respectively the initial and final states. It can be shown * that these two states are both eigenstates of the total Hamiltonian; i.e.,

$$H |n^{in}\rangle = E_n |n^{in}\rangle \tag{6.60}$$

and

$$H |n^{f}\rangle = E_n |n^{f}\rangle . \tag{6.61}$$

For a multiparticle state, $|n^{in}\rangle$ represents plane waves described by (6.56) plus outgoing waves due to the interaction. This is illustrated in Fig. 6.4, in which the free state $|n\rangle$ consists of two particles

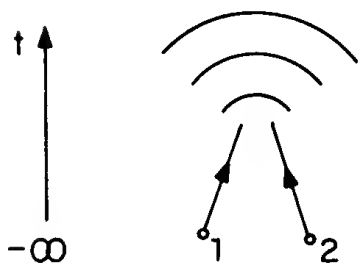


Fig. 6.4. The evolution from a state of free particles 1 and 2 at time = - ∞ to one with outgoing waves at time t .

* See, e.g., S. S. Schweber, An Introduction to Relativistic Quantum Field Theory (New York, Row, Peterson and Co., 1961), pages 320-25, and T. D. Lee and M. Nauenberg, Phys. Rev. 133, B1549 (1964), Appendix A.

1 and 2. Through the time interval from $-\infty$ to t these two particles have interacted continuously, leading to the outgoing-wave component of $U(t, -\infty) \mid n >$. In a similar way, by considering the time evolution from t to $+\infty$, one can show that $\mid n^f >$ represents the superposition of plane waves plus incoming waves. The S-matrix between two states of "free particles" $\mid n >$ and $\mid n' >$ is

$$< n' \mid S \mid n > = < n' \mid U(\infty, -\infty) \mid n >$$

where $\mid n >$ and $\mid n' >$ are both of the form (6.56). Because of (6.58)-(6.59), the matrix element of S can also be written as

$$< n' \mid S \mid n > = < n'^f \mid n^{in} > . \qquad (6.62)$$

Since the sets $\{ \mid n^{in} > \}$ and $\{ \mid n^f > \}$ are each a complete orthonormal set of basis vectors in Hilbert space, the S-matrix is simply the unitary transformation between these two sets. For any scattering process

$$1 + 2 \rightarrow 1' + 2' + \cdots ,$$

the plane-wave part of $\mid n^{in} >$ is associated with the initial particles $1 + 2$, and that of $\mid n^f >$ with the final particles $1' + 2' + \cdots$.

Problem 6.1. Show that to order α^2 the differential and total cross sections of $e^+ + e^- \rightarrow \mu^+ + \mu^-$ are

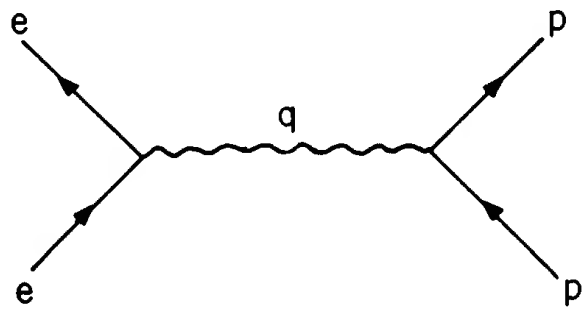
$$d\sigma = \tfrac{1}{2}\pi\alpha^2 v(2 - v^2 \sin^2\theta) \frac{d\cos\theta}{4E^2}$$

and $\sigma = \tfrac{1}{2}\pi\alpha^2 v(1 - \tfrac{1}{3}v^2)/E^2$ where $2E$, v and θ are, respectively, the total energy, the muon velocity and the angle between the e^- and μ^- momenta in the center-of-mass system, $\alpha = e^2/4\pi$ and, for simplicity, the electron mass m_e is set to be 0.

Problem 6.2. Show that to order α^2 the differential cross section of $e^- + p \to e^- + p$, in the approximation that $m_e = 0$ and the strong interaction of the proton is neglected, is

$$\frac{d\sigma}{dq^2} = 4\pi \left(\frac{\alpha}{q^2}\right)^2 \left[1 - \frac{q^2}{q^2_{max}} + \tfrac{1}{2}\left(\frac{q^2}{2m_p E}\right)^2\right]$$

where m_p is the proton mass, E is the initial e^- energy in the rest system of the initial p (laboratory system), and q^2 is the (4-momentum transfer)^2 between e^- and p:



Hence, $q^2 = \vec{q}^2 - q_0^2$ can vary between 0 to

$$q^2_{max} = \frac{2m_p E}{(1 + \frac{m_p}{2E})} .$$

References

J. Bjorken and S. D. Drell, Relativistic Quantum Mechanics (New York, McGraw-Hill, 1964).

R. Feynman, Quantum Electrodynamics (New York, W. A. Benjamin, 1962).

G. Wentzel, Quantum Theory of Fields (New York, Interscience Publishers, Inc., 1949).

Chapter 7

SOLITONS

The usual description of a bound state is in terms of the Schrödinger equation if it is nonrelativistic, or its generalization, the Bethe-Salpeter equation, in the relativistic case. Such an approach is highly successful in the case of atoms and molecules; it is also reasonably adequate with regard to nuclear structure. In these descriptions, Planck's constant plays an essential role. These bound states exist only in quantum mechanics. Indeed, in the case of the coupling between matter and the electromagnetic field, classical physics is totally inadequate to provide a stable atomic structure against radiation. It is that failure which led to the discovery of quantum mechanics in the first place. Since then, it has usually been thought that, in a relativistic field theory, in order to have stationary bound states, quantum mechanics must be crucial. As we shall see, this turns out not to be the case. In a nonlinear field theory, with an appropriate amount of nonlinearity, stable bound states can exist on a classical, as well as quantum mechanical, level. Such bound states are called solitons.

7.1 Early History

The earliest discussion of the subject was given by J. Scott Russell in the Report of the British Association for the Advancement of Science, published in 1845. In his own words (given below):

Report on Waves. By J. Scott Russell, *Esq., M.A., F.R.S. Edin., made to the Meetings in* 1842 *and* 1843.

Members of Committee { Sir John Robison*, *Sec. R.S. Edin.*
J. Scott Russell, *F.R.S. Edin.*

I believe I shall best introduce this phænomenon by describing the circumstances of my own first acquaintance with it. I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped—not so the mass of water in the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well-defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height. Its height gradually diminished, and after a chase of one or two miles I lost it in the windings of the channel. Such, in the month of August 1834, was my first chance interview with that singular and beautiful phænomenon which I have called the Wave of Translation, a name which it now very generally bears; which I have since found to be an important element in almost every case of fluid resistance, and ascertained to be the type of that great moving elevation of the sea, which, with the regularity of a planet, ascends our rivers and rolls along our shores.

Fig. 7.1

Scott Russell then went on to propose that the solitary object which he encountered actually represents a general class of solutions of hydrodynamics, which he first called "wave of translation", and later "solitary wave". Unlike the shock wave, which is singular at the shock front, the "solitary wave" is regular everywhere without singularity. The solitary wave is nondispersive and stable; therefore, it is different from any wave packet composed of the usual plane-wave solutions. However, Scott Russell did not succeed in convincing all his colleagues. As we can see from Fig. 7.2, taken from an 1876 paper by Lord Rayleigh, the subject of the solitary wave was still in hot dispute among various leading physicists of the time. The dispute was not settled until 1895, when Korteweg and de Vries * gave the complete analytic explanation in terms of what is now called the soliton solution of the nonlinear hydrodynamical equation—the Korteweg-de Vries equation.

Nevertheless, the question remains whether such stable, nonsingular and nondispersive solutions can occur in other domains of physics, outside hydrodynamics. This problem received a new impetus through the work done by Fermi, Pasta and Ulam ** in the early fifties. By using one of the first large electronic computers, Maniac I, they investigated the approach to equipartition of energy between 64 harmonic oscillators, coupled with some very weak nonlinear couplings. Initially, all energy lay only in one oscillator. To their great surprise, the usual idea of how the thermal equilibrium is reached turned out to be quite incorrect.

---

* D. J. Korteweg and G. de Vries, Phil. Mag. 39, 422 (1895).

** Collected Papers of Enrico Fermi, general editor E. Segré (University of Chicago Press, 1965), Vol. II, 978.

THE

LONDON, EDINBURGH, AND DUBLIN

PHILOSOPHICAL MAGAZINE

AND

JOURNAL OF SCIENCE.

[FIFTH SERIES.]

APRIL 1876.

XXXII. *On Waves.* *By* LORD RAYLEIGH, *M.A., F.R.S.**

. . .

The Solitary Wave.

This is the name given by Mr. Scott Russell to a peculiar wave described by him in the British-Association Report for 1844.

. . .

Airy, in his treatise on Tides and Waves, still probably the best authority on the subject, appears not to recognize any thing distinctive in the solitary wave.

. . .

On the other hand, Professor Stokes says* :—"It is the opinion of Mr. Russell that the solitary wave is a phenomenon *sui generis*, in no wise deriving its character from the circumstances of the generation of the wave.

Fig. 7.2

266.

STUDIES OF NON LINEAR PROBLEMS

E. FERMI, J. PASTA, and S. ULAM
Document LA-1940 (May 1955).

ABSTRACT.

A one-dimensional dynamical system of 64 particles with forces between neighbors containing nonlinear terms has been studied on the Los Alamos computer MANIAC I. The nonlinear terms considered are quadratic, cubic, and broken linear types. The results are analyzed into Fourier components and plotted as a function of time.

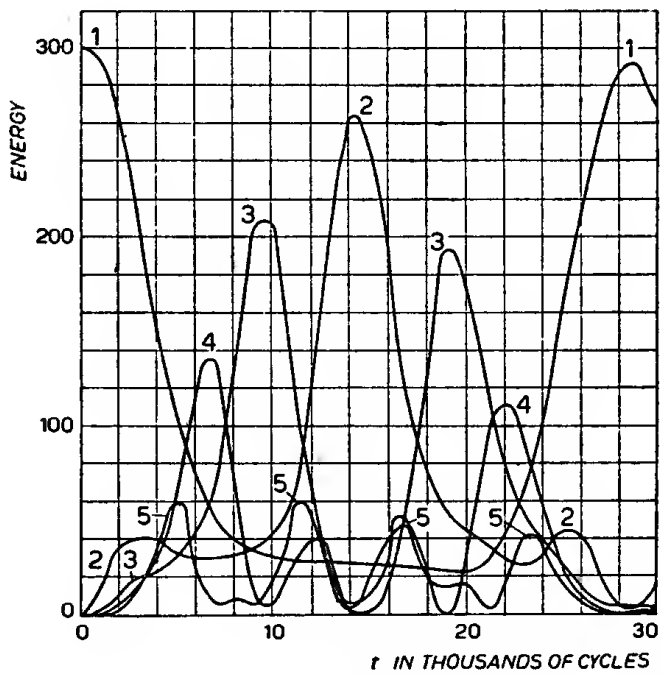


Fig. 1. – The quantity plotted is the energy (kinetic plus potential in each of the first five modes). The units for energy are arbitrary. $N = 32$; $\alpha = 1/4$; $\delta t^2 = 1/8$. The initial form of the string was a single sine wave. The higher modes never exceeded in energy 20 of our units. About 30,000 computation cycles were calculated.

Fig. 7.3

As we can see from Fig. 7.3, after some tens of thousands of cycles, the energy invariably returned nearly completely to the original mode, leaving only a few percent of the total energy to very few other oscillators. [This is not the Poincaré cycle, which requires a much longer time duration.] The development of such collective modes is a general phenomenon; it can be approximately represented by the soliton solution of the Toda lattice.* An important and general feature of the soliton solutions is that they exist even if the nonlinear coupling is extremely weak:

$$\text{weak coupling} \neq \text{weak amplitude}. \tag{7.1}$$

Since then, there has been a large number of papers on soliton solutions. The review article by Scott, Chu and McLaughlin** in 1973 listed a total of 267 references. However, all of these dealt only with classical soliton solutions, and almost all were restricted to one space-dimension and to only seven specific equations: Korteweg-de Vries equation, sine-Gordon equation, etc. Recently, there has been some major progress made in this field, both in classical solutions, extending them to three space-dimensions, and in quantum soliton solutions, developing general techniques so that (at least for boson fields in the weak-coupling limit) to each classical soliton solution, there exists a corresponding quantum solution. These new developments will be the main part of our discussion.

---

* M. Toda, Progr. Theor. Phys. Suppl. 45, 174 (1970).

** A. C. Scott, F. Y. F. Chu and D. W. Mclaughlin, Proc. IEEE 61, 1443 (1973).

7.2 Definition, Classification and Some General Remarks

Let us begin with the definition:

A classical soliton is any spatially confined and nondispersive solution of a classical field theory.

Throughout our discussions we shall be interested only in relativistic local field theories. In order to have soliton solutions, there must be nonlinear couplings; otherwise, the only solutions are plane waves. While wave packets can be formed through the superposition of plane waves, these packets are always dispersive and therefore not solitons.

The following remarks are applicable to any boson-field solitons.

1. In a general case, the Lagrangian may consist of several fields and many different couplings. It is convenient to represent the various fields collectively as ϕ, and to write the Lagrangian density $\mathcal{L}$ as

$$\mathcal{L} = -\tfrac{1}{2}\left(\frac{\partial\phi}{\partial x_\mu}\right)^2 - \frac{1}{g^2}\, V(g\phi) \tag{7.2}$$

where g is dimensionless and V has its minimum at $\phi = 0$. Without any loss of generality, this minimum value may be chosen to be zero. Thus, in a power series expansion

$$\frac{1}{g^2}\, V(g\phi) = \tfrac{1}{2} m^2 \phi^2 + O(g\phi^3) + O(g^2\phi^4) + \cdots, \tag{7.3}$$

in which the quadratic ϕ^2-term is independent of g. If there is only a single field in the theory, then m is a number. Otherwise, ϕ represents a column matrix with n components, and $\frac{1}{2}m^2\phi^2$ stands for $\frac{1}{2}\tilde{\phi}\, m^2 \phi$ where m is an $n \times n$ matrix. The equation of motion can be obtained through the variational principle (2.10); it is

$$\frac{\partial^2\phi}{\partial x_\mu^2} - \frac{1}{g}\, V'(g\phi) = 0 \tag{7.4}$$

where $V'(\sigma) = dV(\sigma)/d\sigma$ and $\sigma = g\phi$. Because of (7.3), the above equation becomes

$$\frac{\partial^2\phi}{\partial x_\mu^2} = m^2\phi + O(g\phi^2) + O(g^2\phi^3) + \cdots . \tag{7.5}$$

When $g = 0$ the equation becomes linear. Therefore, g characterizes the various nonlinear couplings in the equation. As mentioned before, since soliton solutions are nondispersive, they do not exist when $g = 0$. As we shall see, all soliton solutions are singular when $g \to 0$.

2. In a classical theory, this singularity is always a simple pole. To show this, we may write

$$\phi_{classical} = \frac{1}{g}\sigma . \tag{7.6}$$

The Lagrangian density $\mathcal{L}$ then becomes

$$\mathcal{L} = \frac{1}{g^2}\mathcal{L}_\sigma \tag{7.7}$$

where

$$\mathcal{L}_\sigma = -\tfrac{1}{2}\left(\frac{\partial\sigma}{\partial x_\mu}\right)^2 - V(\sigma) \tag{7.8}$$

which is g-independent. Since the classical solution is determined by the extremity of the action integral, the g-independence of $\mathcal{L}_\sigma$ implies that the corresponding soliton solution σ is also g-independent, and that establishes (7.6). Therefore, the existence of soliton solutions does not depend on the strength of g, so long as $g \neq 0$. This is why, as noted in (7.1), even in the case of weak coupling it is not possible to neglect the soliton solutions. Unlike the plane-wave solution, the soliton solution $\to \infty$ when $g \to 0$.

3. An important and delightful feature which was discovered only relatively recently is that for any boson-field system, once the

classical soliton exists there is always a corresponding quantum soliton solution, at least in the weak coupling.

The simplest way to anticipate this is to note that the action A in a quantum theory can be set to be $\hbar^{-1}$ times the classical action; i.e.

$$A = \hbar^{-1} \int \mathcal{L}\, d^4x = (\hbar g^2)^{-1} \int \mathcal{L}_\sigma\, d^4x \qquad (7.9)$$

where $\mathcal{L}$ and $\mathcal{L}_\sigma$ are related by (7.7). In the quantum theory, one considers all paths leading from an initial to a final configuration. [See Fig. 7.4.] Each path carries an amplitude proportional to e^{iA},

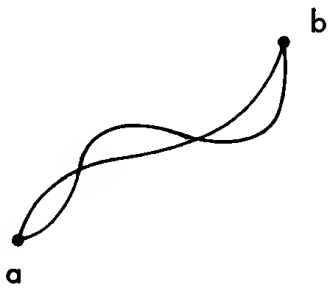


Fig. 7.4. In quantum theory each path from a to b carries an amplitude proportional to e^{iA} where A is the action integral. [See Chapter 19.]

and the superposition of all these amplitudes is the state vector. When $\hbar$ approaches 0, the only important path is the one with a stationary phase, $\delta A = 0$, and that leads to the classical description. The formal expansion in terms of $\hbar$ gives the familiar W. K. B. approximation. Because of (7.9), we expect g^2 to play the same role as $\hbar$:

$$g^2 \sim \hbar \,. \qquad (7.10)$$

Therefore, an expansion in g^2 is equivalent to that in $\hbar$; the leading term must be the same as the classical limit. The details of how to carry out such a quantum expansion will be given in Section 7.6. Here, we only note that (7.10) explains why in the weak coupling the existence of a classical soliton implies a corresponding quantum solution. For example, on account of (7.6)-(7.7), the energy of a classical soliton is of the form

$$E_{classical} = O(g^{-2}) \quad . \tag{7.11}$$

In a perturbation expansion, the energy of the corresponding quantum solution becomes then

$$E_{quantum} = E_{classical}\,[1 + O(g^2) + O(g^4) + \cdots] \quad . \tag{7.12}$$

Thus,

$$E_{quantum} \to E_{classical} \qquad \text{when} \quad g \to 0 \quad , \tag{7.13}$$

at least formally.

Another pleasant aspect is that within the conventional class of renormalizable theories, all radiative corrections $O(g^2)$, $O(g^4)$, $\cdots$ in (7.12) are expected to be automatically finite for the soliton solutions. This is closely tied to the fact that the classical soliton solution is regular everywhere. At very high frequencies, the scattering amplitude of an incident plane wave by the soliton must be negligibly small; hence, the existence of soliton solutions should not alter the high-energy behavior of the theory. For the renormalizable theories, because all radiative corrections are finite, the limit (7.13) is valid in a real sense, and thereby connects the classical to the quantum solution.

Because of the uncertainty principle, a quantum soliton cannot be confined in space all the time. The definition of a quantum soliton

is tied to that of the corresponding classical solution through (7.12)-(7.13).

If we restrict ourselves to renormalizable relativistic local field theories, then all soliton solutions can be classified into two general types (the details of which will be given in the next few sections).

(1) Topological solitons. The necessary condition is that there should be degenerate vacuum states so that the boundary condition at infinity for a soliton state is topologically different from that of the physical vacuum state. Some typical examples of the topological soliton solutions are those of the sine-Gordon equation* in one space-dimension, the vortex solution of Nielsen and Olesen** in two space-dimensions, and the magnetic monopole solution of 't Hooft and Polyakov*** in three space-dimensions.

(2) Nontopological solitons. The boundary condition at infinity for a nontopological soliton is the same as that for the vacuum state. Thus, there is no need of the degenerate vacuum states. The necessary condition for the existence of nontopological solitons is that there should be an additive conservation law. The nontopological soliton solutions can also exist in any space dimension****, as we shall discuss in the subsequent sections.

---

* See Problem 7.1 for its definition.

** H. B. Nielsen and P. Olesen, Nucl. Phys. B61, 45 (1973).

*** G. 't Hooft, Nucl. Phys. B79, 276 (1974); A. M. Polyakov, JETP Lett. 20, 194 (1974).

**** R. Friedberg, T. D. Lee and A. Sirlin, Phys. Rev. D13, 2739 (1976), Nucl. Phys. B115, 1, 32 (1976).

7.3 One-space-dimensional Examples

For simplicity, we first consider soliton solutions in one space-dimension (plus the time-dimension). In view of (7.10)-(7.13), we need only examine the classical system. The quantum solution can then be derived by the perturbation series, as will be shown in Section 7.6.

1. Topological soliton. Let ϕ be a Hermitian field. In accordance with (7.2), the Lagrangian density can be written as

$$\mathcal{L} = -\tfrac{1}{2}\left(\frac{\partial\phi}{\partial x_\mu}\right)^2 - \frac{1}{g^2}\,V(g\phi) \quad , \tag{7.14}$$

where $x_\mu = (x, it)$. Through the substitution

$$\phi = \frac{1}{g}\,\sigma \quad ,$$

$\mathcal{L}$ becomes

$$\mathcal{L} = \frac{1}{g^2}\,\mathcal{L}_\sigma \tag{7.15}$$

where

$$\mathcal{L}_\sigma = -\tfrac{1}{2}\left(\frac{\partial\sigma}{\partial x_\mu}\right)^2 - V(\sigma) \quad . \tag{7.16}$$

Since the necessary condition for the topological soliton is the existence of degenerate vacuum, there must be more than one minimum of $V(\sigma)$. Without any loss of generality, we may choose the minimum of V to be 0. Therefore, as shown in Fig. 7.5, V is ≥ 0 and it has more than one minimum, say $V(\sigma) = 0$, at $\sigma = a, b, \cdots$. In terms of σ, the equation of motion is

$$\frac{\partial^2\sigma}{\partial x_\mu^2} - \frac{dV}{d\sigma} = 0 \tag{7.17}$$

whose time-independent solution is determined by

$$\frac{d^2\sigma}{dx^2} - \frac{dV}{d\sigma} = 0 \quad .$$

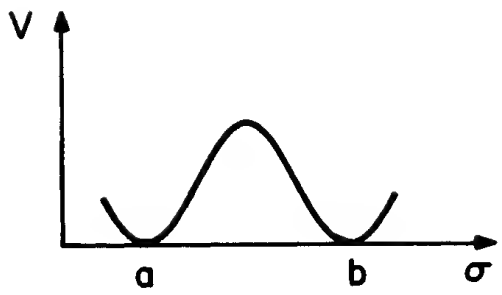


Fig. 7.5. The schematic drawing of $V(\sigma)$ for a theory that has topological soliton solutions.

We may multiply it by $d\sigma/dx$ and then integrate. This leads to

$$\tfrac{1}{2}\left(\frac{d\sigma}{dx}\right)^2 - V(\sigma) = \text{constant} \quad . \tag{7.18}$$

There exists a simple mechanical analog: We may consider a point particle with σ as its "position" and x its "time", moving in a "potential" $-V(\sigma)$, as shown in Fig. 7.6. The above equation is then simply the energy conservation law in the analog problem. In this analog problem, let us set at "time" $x = -\infty$ the "position" of the particle σ at a. We may start the motion by pushing the

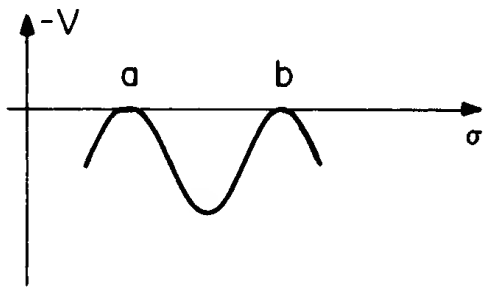


Fig. 7.6. The "potential" $-V$ in the mechanical analog problem.

particle very gently towards the right. As the "time" x increases, σ moves from a towards b; as $x \to \infty$, $\sigma \to b$, because of energy conservation. Its analytic expression is

$$x = \int^{\sigma} \frac{d\sigma'}{\sqrt{2V(\sigma')}} \tag{7.19}$$

which contains an integration constant ξ, as shown in Fig. 7.7.

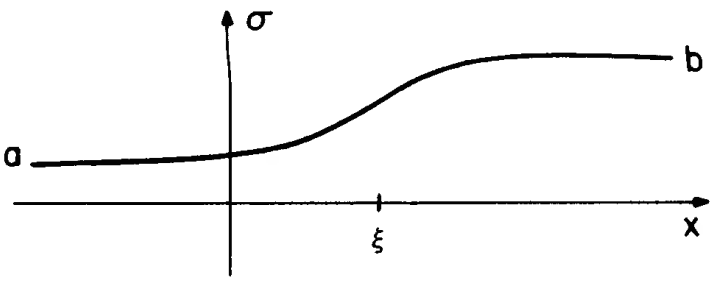


Fig. 7.7. A topological soliton solution.

Returning now to the original field-theory problem, because of (7.14)-(7.16), the energy density for a time-independent solution can be written as

$$\mathcal{E}(x) = \frac{1}{g^2}\,\mathcal{E}_{\sigma}(x)$$

where

$$\mathcal{E}_{\sigma}(x) = \tfrac{1}{2}\left(\frac{d\sigma}{dx}\right)^2 + V(\sigma) \quad . \tag{7.20}$$

From Fig. 7.7 and the fact that $V(a) = V(b) = 0$ we find that the energy density $\mathcal{E}(x)$ is of the form given in Fig. 7.8, which is confined in space at all times. Because the boundary conditions of the field $\phi = \sigma/g$ at $x = \pm\infty$ are different, it is called the topological

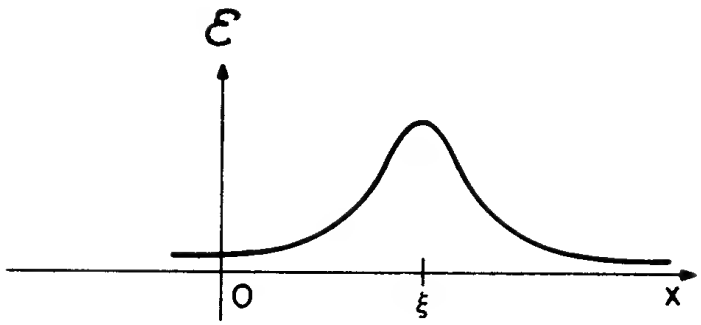


Fig. 7.8. A schematic drawing of energy density for the soliton given by Fig. 7.7.

soliton solution. Its stability is insured by the boundary conditions at infinity.

In the mechanical-analog problem, we may derive another solution by setting at "time" $x = -\infty$, the "position" σ at b. As x increases, σ moves from b to a. If we call the solution given by Fig. 7.7 the soliton, then this new solution is the anti-soliton. Both have the same energy. Thus, the concept of particle-antiparticle conjugation already exists on the classical level.

Because of Lorentz invariance, if $\sigma(x)$ is a solution of (7.17), then

$$\sigma(\gamma x - \gamma v t)$$

must also satisfy the same field equation, where $\gamma = (1 - v^2)^{-\frac{1}{2}}$. Consequently, we also have the solution for a moving soliton, or anti-soliton.

To describe the scattering between a soliton and an anti-soliton (or between two solitons or anti-solitons) we may consider the initial condition that at $t = -\infty$, one of them is, e.g., at $x = -\infty$

moving with velocity $v > 0$, while the other is at $x = +\infty$ moving with velocity $-v$. In general, the state will change in the course of time due to collision. In the special case of the sine-Gordon equation (defined in Problem 7.1), because of the presence of an infinite number of conservation laws, the shape and velocity of each soliton or anti-soliton remain unchanged even after such a head-on collision. We refer to this special class as indestructible solitons. Indestructible solitons exist only in one space-dimension. If one requires relativistic invariance, then it exists only for the sine-Gordon equation.

2. Nontopological soliton

To construct nontopological solitons, one does not need degenerate vacuum. However, as we shall see, because of the requirement of an additive conservation law, there must at least be a complex field.

Again, we shall first consider the case of one space-dimension. Let ϕ be a complex field. In accordance with our general form (7.2), the Lagrangian density is assumed to be

$$\mathcal{L} = -\frac{\partial \phi^\dagger}{\partial x_\mu}\frac{\partial \phi}{\partial x_\mu} - \frac{1}{g^2}\,U(g^2\phi^\dagger\phi) \tag{7.21}$$

where † denotes the Hermitian conjugate. By using the variation principle (2.10), one finds the equation of motion to be

$$\frac{\partial^2 \phi}{\partial x_\mu^2} - \phi\,\frac{d}{d(g^2\phi^\dagger\phi)}\,U(g^2\phi^\dagger\phi) = 0 \tag{7.22}$$

where $x_\mu = (x, it)$, as before. We shall assume $U(\phi^\dagger\phi)$ has a single minimum at $\phi = 0$. Furthermore, the minimum value of U is 0. Hence, just as in (7.3),

$$\frac{1}{g^2}\,U \to m^2\phi^\dagger\phi \qquad \text{as} \qquad \phi \to 0 \tag{7.23}$$

where m is the mass of the usual plane wave solution. [See (7.35)-(7.36) below.]

The Lagrangian density (7.21) is invariant under the phase transformation

$$\phi \rightarrow e^{-i\theta}\phi \quad . \tag{7.24}$$

Hence, as can be verified directly, the current

$$j_\mu = i\phi^\dagger \frac{\partial \mathcal{L}}{\partial(\frac{\partial \phi^\dagger}{\partial x_\mu})} - i \frac{\partial \mathcal{L}}{\partial(\frac{\partial \phi}{\partial x_\mu})}\phi \tag{7.25}$$

satisfies

$$\frac{\partial j_\mu}{\partial x_\mu} = 0 \quad . \tag{7.26}$$

On account of (7.21), in our case the current j_μ is given by

$$j_\mu = i \frac{\partial \phi^\dagger}{\partial x_\mu}\phi - i\phi^\dagger \frac{\partial \phi}{\partial x_\mu} \quad . \tag{7.27}$$

The particle density ρ is given by the time-component of j_μ multiplied by $-i$. From (7.27), one finds

$$\rho = i(\phi^\dagger \dot{\phi} - \dot{\phi}^\dagger \phi) \quad . \tag{7.28}$$

Its space integral is the particle number N,

$$N = \int \rho \, dx \quad . \tag{7.29}$$

Because of (7.26) N is conserved; i.e.

$$\dot{N} = 0 \quad . \tag{7.30}$$

From (7.28)-(7.29) one sees that for $N \neq 0$, ϕ must vary with time. It is not difficult to show that the lowest-energy classical solution should be of the form

$$\phi = \frac{1}{g}\sigma(x)\, e^{-i\omega t} \tag{7.31}$$

where $\sigma(x)$ is real. In terms of σ, (7.22) becomes

$$\frac{d^2\sigma}{dx^2} + \omega^2 \sigma - \sigma \frac{d}{d\sigma^2} U(\sigma^2) = 0 \quad , \tag{7.32}$$

which, after being multiplied by $d\sigma/dx$, can be integrated. The result is

$$\tfrac{1}{2} \left(\frac{d\sigma}{dx}\right)^2 - V(\sigma) = \text{constant} \tag{7.33}$$

where

$$V(\sigma) = \tfrac{1}{2} U(\sigma^2) - \tfrac{1}{2} \omega^2 \sigma^2 \quad . \tag{7.34}$$

In Fig. 7.9, an example of the function U is plotted against σ ; as mentioned before, U has a single minimum at $\sigma = 0$.

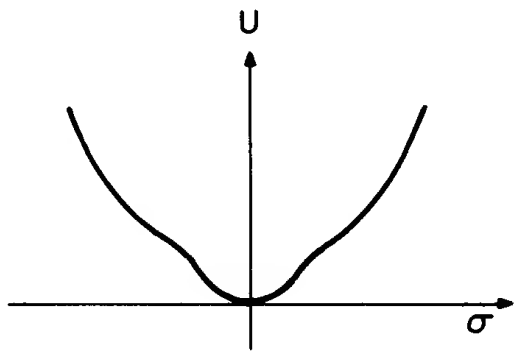


Fig. 7.9. An example of $U(\sigma^2)$ vs. σ .

Let Ω be the volume that encloses the whole system. When $\Omega \to \infty$, (7.22) admits the usual plane-wave solutions

$$\phi = \sqrt{\frac{N}{2\omega\Omega}} \, e^{i(kx - \omega t)} \tag{7.35}$$

where

$$\omega = \sqrt{\vec{k}^2 + m^2} \quad . \tag{7.36}$$

This is because in this limit the amplitude ϕ becomes infinitesimal;

therefore, on account of (7.23), (7.22) reduces to

$$\frac{\partial^2 \phi}{\partial x_\mu^2} - m^2 \phi = 0 \qquad (7.37)$$

and (7.35) is the solution. The soliton solution differs from the plane wave solution, since at finite x its amplitude does not become infinitesimal as $\Omega \to \infty$. Furthermore, when $x \to \pm \infty$, the soliton amplitude approaches zero exponentially; therefore $\omega^2 < m^2$. Hence we may regard these two types of solution as analytical continuations of each other:

$\omega^2 > m^2$ for the plane-wave solution,

and $\omega^2 < m^2$ for the soliton solution. (7.38)

This relation is valid in any space-dimension.

We shall now show that in order to have the nontopological soliton solutions, the function $V = \frac{1}{2}(U - \omega^2 \sigma^2)$, defined by (7.34), must be of the form given by Fig. 7.10, at least when $\omega^2 = m^2$. More specifically, the condition

$$U(\sigma^2) - \omega^2 \sigma^2 = 0 \qquad (7.39)$$

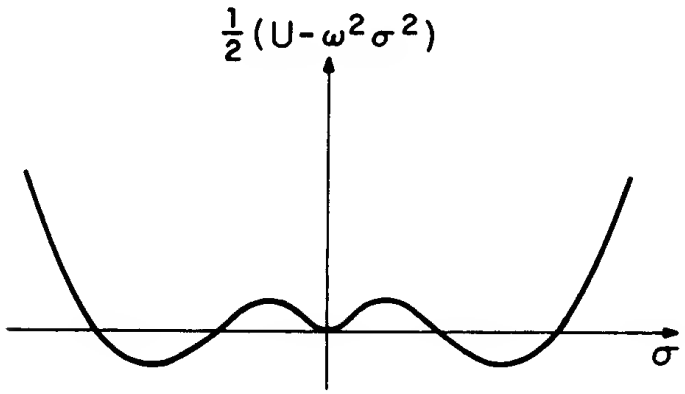


Fig. 7.10. A schematic drawing of $V = \frac{1}{2}(U - \omega^2 \sigma^2)$.

has, for $\omega^2 \leq m^2$, besides the solution $\sigma = 0$ also some other $\sigma \neq 0$ solutions. Assuming that this is indeed the case, just as in the previous example of the topological soliton, we may consider the mechanical analog in which there is a point particle at a "position" σ and a "time" x, moving in a potential $-V = -\frac{1}{2}[U - \omega^2\sigma^2]$, shown in Fig. 7.11. At the "time" $x = -\infty$, we may set the particle at the

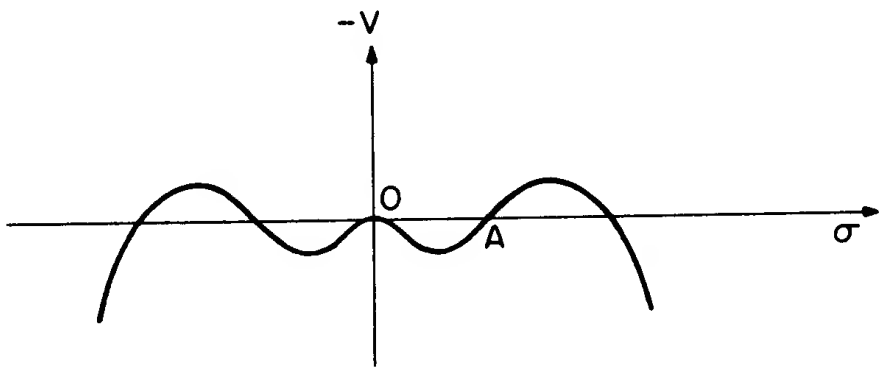


Fig. 7.11. The "potential" $-V$ in the mechanical analog problem.

"position" $\sigma = 0$. Again, we may start the motion by an extremely gentle push towards the right. As x increases, σ moves to A and then returns to 0 at $x = +\infty$. The general solution is given by

$$x - \xi = \int_A^\sigma \frac{d\sigma}{\sqrt{2V(\sigma)}} \tag{7.40}$$

where ξ is the integration constant. A schematic drawing of the solution is given in Fig. 7.12. When $x = \xi$, $\sigma = A$. At both infinities, the nontopological soliton solution σ satisfies the same boundary condition:

$$\sigma \to 0 \quad \text{when} \quad x \to \infty \quad \text{or} \quad -\infty .$$

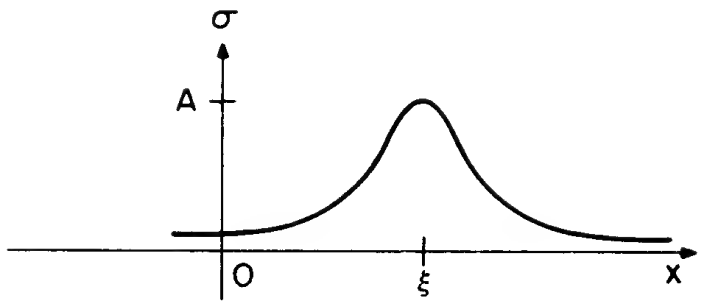


Fig. 7.12. A nontopological soliton solution.

We note that when $\sigma \to 0$, on account of (7.23), the function $V = \frac{1}{2}(U - \omega^2\sigma^2)$ becomes

$$V \to \tfrac{1}{2}(m^2 - \omega^2)\,\sigma^2 + O(\sigma^4) \quad . \tag{7.41}$$

Thus, in order that in Fig. 7.10 the curve $V(\sigma)$ should be concave upward at the origin $\sigma = 0$, we must have

$$\omega^2 < m^2 \quad ,$$

which confirms (7.38). Furthermore, condition (7.39) can be most easily satisfied if the σ^4-term is < 0, which corresponds to attraction between the fields.

As an explicit example, we may refer to Problem 7.2 in which

$$\frac{1}{g^2}\,U = \frac{m^2\phi^\dagger\phi}{1+\epsilon^2}\,[(1 - g^2\phi^\dagger\phi)^2 + \epsilon^2] \quad .$$

The solution is

$$\phi = \frac{1}{g}\left[\frac{a}{1 + \sqrt{1-a}\,\cosh y}\right]^{\frac{1}{2}} e^{-i\omega t} \tag{7.42}$$

where

$$a = (1+\epsilon^2)\,\frac{1}{m^2}\,(m^2 - \omega^2)$$

and

$$y = 2\sqrt{m^2 - \omega^2}\ (x - \xi) \ .$$

Equations (7.19) and (7.40) reduce the problem of finding any one-space-dimensional soliton solution, topological or nontopological, to quadrature.

7.4 Derrick Theorem

A theorem due to G. H. Derrick* imposes severe restrictions on the types of soliton solutions that can exist when the space-dimension (excluding the time-dimension) is $D > 1$. Let us consider a classical system consisting only of scalar fields $\phi_1, \cdots, \phi_N$, whose Lagrangian density is

$$\mathcal{L} = \tfrac{1}{2} \sum_a [\, \dot{\phi}_a^2 - (\vec{\nabla}\phi_a)^2 \,] - U(\phi_a)$$

where $a = 1, 2, \cdots, N$, $\vec{\nabla}$ is the D-dimensional gradient vector and U is assumed to be $\geqslant 0$, with its minimum value given by

$$\min\ U(\phi_a) = 0 \ . \tag{7.43}$$

Thus, the ground state (i.e., the vacuum) is of zero energy. There may, however, exist more than one such ground state.

Theorem. For $D \geqq 2$, the only time-independent solutions of finite energy are the ground states; i.e., ϕ_a = constant everywhere for which $U(\phi_a) = 0$.

Proof. From the Lagrangian density, it follows that the Hamiltonian density is

$$\mathcal{H} = \tfrac{1}{2} \sum_a [\, \pi_a^2 + (\vec{\nabla}\phi_a)^2 \,] + U(\phi_a)$$

---

* G. H. Derrick, J. Math. Phys. 5, 1252 (1964).

where $\Pi_a = \dot{\phi}_a$ is the conjugate momentum of ϕ_a. Let $\phi_a(\vec{r})$ be a time-independent solution of the theory; hence $\phi_a(\vec{r})$ satisfies the field equation

$$\nabla^2 \phi_a - \frac{dU}{d\phi_a} = 0 \quad .$$

The corresponding total energy $\int \mathcal{H} d^3r$ is

$$E(1) = T_1 + V_1$$

with

$$T_1 \equiv \tfrac{1}{2} \sum_a \int (\vec{\nabla} \phi_a(\vec{r}))^2 \, d^D r$$

and

$$V_1 \equiv \int U(\phi_a(\vec{r})) \, d^D r \quad .$$

We may construct a new function $\phi_a^{\lambda}(\vec{r})$, defined by

$$\phi_a^{\lambda}(\vec{r}) \equiv \phi_a(\lambda \vec{r}) \quad .$$

If everywhere the field distribution is set to be this new function $\phi_a^{\lambda}(\vec{r})$, since it is also time-independent the corresponding total energy is now given by

$$E(\lambda) = T_\lambda + V_\lambda$$

where

$$T_\lambda = \tfrac{1}{2} \sum_a \int (\vec{\nabla} \phi_a^{\lambda}(\vec{r}))^2 \, d^D r$$

and

$$V_\lambda \equiv \int U(\phi_a^{\lambda}(\vec{r})) \, d^D r \quad .$$

Because $\phi_a^{\lambda}(\vec{r})$ is obtained from $\phi_a(\vec{r})$ through the scale transformation

$$\vec{r} \rightarrow \lambda \vec{r} \quad ,$$

one sees that T_λ and V_λ are related to their values at $\lambda = 1$ by

$$T_\lambda = \lambda^{2-D} T_1$$

and

$$V_\lambda = \lambda^{-D} V_1 \quad .$$

Hence, the derivative of $E(\lambda)$ with respect to $\ln \lambda$ is

$$\lambda \frac{dE(\lambda)}{d\lambda} = (2 - D)\, T_\lambda - D V_\lambda \quad .$$

When $\lambda = 1$, $\phi_a^\lambda(\vec{r})$ becomes $\phi_a(\vec{r})$, which is a solution of the field equation, and therefore its total energy must be stationary against any small variations. Consequently $dE(\lambda)/d\lambda$ is zero at $\lambda = 1$, from which we obtain

$$(2 - D)\, T_1 - D\, V_1 = 0 \quad . \tag{7.44}$$

Now, for $D \geqq 2$, both $(2 - D)\, T_1$ and $- D\, V_1$ are $\leqq 0$. If D is > 2, (7.44) is clearly impossible unless $T_1 = V_1 = 0$. For $D = 2$, (7.44) implies $V_1 = 0$, which means that $U(\phi_a) = 0$ everywhere and $E(1)$ equals T_1. Suppose that $U(\phi_a) = 0$ when $\phi_a =$ either constant: c_a or c_a'. If $\phi_a = c_a$ in one region of space and $\phi_a = c_a'$ in a neighboring region, then $\vec{\nabla}\phi_a$ across the boundary of these two regions contains a δ function and the integral of $(\vec{\nabla}\phi_a)^2$ gives ∞. Therefore, the only finite-energy solution is $\phi_a =$ same constant everywhere. This then completes the proof.

Thus, when D is > 1, in order to have soliton solutions, we must either include fields of nonzero spin, or consider time-dependent but nondispersive solutions. The former leads to the gauge-field topological solitons, such as the aforementioned vortex solution of Nielsen and Olesen in two dimensions and the magnetic monopole solution of 't Hooft and Polyakov in three dimensions. The latter is represented by the multidimensional nontopological solitons whose properties will be discussed in the next section.

7.5 Solitons vs. Plane Waves

As already noted in (7.35), any nonlinear relativistic field equation always admits plane wave solutions of the form

$$\sqrt{\frac{N}{2\omega\Omega}}\; e^{i\vec{k}\cdot\vec{r} - i\omega t} \tag{7.45}$$

where $\Omega \to \infty$ is the volume of the system and $\omega = \sqrt{\vec{k}^2 + m^2}$.

For the topological soliton, its stability is insured against decay into plane waves because of the different boundary conditions satisfied by these two different types of solutions. For the nontopological soliton, its stability depends on which type of solution is of the lowest energy. In the following, this question will be analysed in some detail.

Let us consider a nonlinear theory in which, say, the particle number N is conserved; furthermore, we assume that the theory has nontopological soliton solutions. For the plane-wave solution (7.45), the energy is linear in N. Hence

$$E\,(\text{plane wave}) = N\omega \quad . \tag{7.46}$$

Since

$$\omega\,(\text{plane wave}) \geqq m \quad ,$$

we have

$$E\,(\text{plane wave}) \geqq Nm \quad . \tag{7.47}$$

For the soliton solutions, the energy E is a nonlinear function of N, and as noted in (7.38) the corresponding ω is $< m$.

1. <u>One space-dimension</u>

As we shall prove, in one space-dimension, if the nontopological solution exists, then its lowest energy E is always lower than Nm:

$$E\,(\text{soliton}) < Nm \quad , \tag{7.48}$$

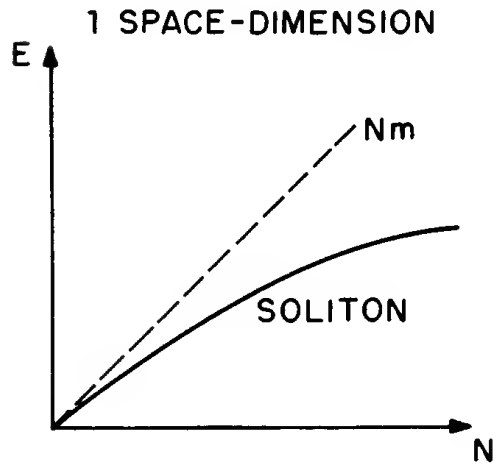


Fig. 7.13. At a given N, the energy of a one-space-dimensional soliton is always lower than that of a plane-wave solution.

which is illustrated in Fig. 7.13. To show (7.48) we first note that N and the phase angle $\theta \equiv \omega t$ of the complex field ϕ are conjugate variables. From Hamilton's equations we have

$$\dot{N} = - \frac{\partial H}{\partial \theta} \tag{7.49}$$

and

$$\dot{\theta} = \frac{\partial H}{\partial N} \ . \tag{7.50}$$

Due to the invariance under the phase transformation (7.24), the Hamiltonian H is independent of θ. Therefore (7.49) gives (7.30); i.e., N is conserved. Since $\omega = \dot{\theta}$ and the value of H is the energy E, (7.50) may be written as

$$\omega = \frac{dE}{dN} \ . \tag{7.51}$$

Now, according to (7.38)

$$\omega\,(\text{soliton}) < m \quad , \tag{7.52}$$

(7.51) implies

$$\frac{d}{dN}\, E\,(\text{soliton}) = \omega < m \quad . \tag{7.53}$$

Next, we shall show that as $\omega \to m-$ and at large $|x|$, the amplitude of the nontopological soliton is of order

$$|\phi| \sim \frac{\sqrt{m^2-\omega^2}}{g}\, e^{-\sqrt{m^2-\omega^2}\,|x|} \quad . \tag{7.54}$$

As can be seen from (7.42), this asymptotic behavior holds at least for the example cited. That this is true in general follows from the fact that when $|x|$ is large, $\phi = \frac{\sigma}{g}$ must be small. From (7.41), we have

$$V \to \tfrac{1}{2}(m^2-\omega^2)\,\sigma^2 + O(\sigma^4) \quad , \qquad \text{as} \qquad \sigma \to 0 \quad .$$

In (7.54), the exponential factor arises from equating $-\frac{d^2\sigma}{dx^2}$ with $\frac{dV}{d\sigma}$; the multiplicative factor $\frac{\sqrt{m^2-\omega^2}}{g}$ comes from the approximate equipartition between the quadratic and the quartic terms in $V(\sigma)$; i.e.

$$(m^2-\omega^2)\,\sigma^2 \sim \sigma^4 \quad . \tag{7.55}$$

Thus, (7.54) is valid in any space-dimension. In one space-dimension, as $\omega \to m-$

$$N \to 2m \int |\phi|^2\, dx \sim \frac{1}{g^2}\sqrt{m^2-\omega^2} \to 0 \quad , \tag{7.56}$$

since $|\phi|^2$ carries a factor $\frac{(m^2-\omega^2)}{g^2}$ and the x-integration gives another factor $\frac{1}{\sqrt{m^2-\omega^2}}$. In Figure 7.13 the dashed line represents the lowest-energy state $\omega = m$ of the plane-wave solution. Since the nontopological soliton solution is the analytic continuation

of the plane-wave solution to the region $\omega < m$, the curve E(soliton) vs. N should be connected to the straight line E(plane wave) = Nm, when $\omega \to m-$. From (7.56), this connection occurs at $N = 0$. Hence (7.53) leads to, for the one-space-dimensional nontopological soliton,

$$E(\text{soliton}) = \int_0^N \omega \, dN < m \int_0^N dN = mN \ , \qquad (7.57)$$

which establishes (7.48) for any N and g.

2. Two space-dimensions

Generalization of the nontopological soliton solutions to space-dimensions > 1 has been given in the literature. A detailed discussion lies outside the scope of this book. As in the one-space-dimensional case, all one needs is some nonlinear interaction which gives rise to attraction between the fields (or field quanta); this would be the case if the interaction is mediated by a scalar field. By using variational considerations, it is not difficult to show that

$$E(\text{soliton}) < Nm \ , \qquad \text{as} \quad N \to \infty \ . \qquad (7.58)$$

In either the two- or three-space-dimensional cases, by following the same argument which led to (7.54) we see that, as $\omega \to m-$ and at large radial distance r, the amplitude of the nontopological soliton is of order

$$|\phi| \sim \frac{\sqrt{m^2 - \omega^2}}{g} \, e^{-\sqrt{m^2 - \omega^2}\, r} \ . \qquad (7.59)$$

Thus, in two space-dimensions, as $\omega \to m-$

$$N \to 2m \int |\phi|^2 \, d^2 r = N_c \sim O(1) \ . \qquad (7.60)$$

This is because the factor $(m^2-\omega^2)$ in $|\phi|^2$ is exactly cancelled by a corresponding factor $(m^2-\omega^2)^{-1}$ in the d^2r integration. The result is given in Fig. 7.14. There exists a critical number N_c. The

nontopological soliton solution exists only for $N > N_c$; in that region, the lowest-energy solution is always the soliton, not the plane wave.

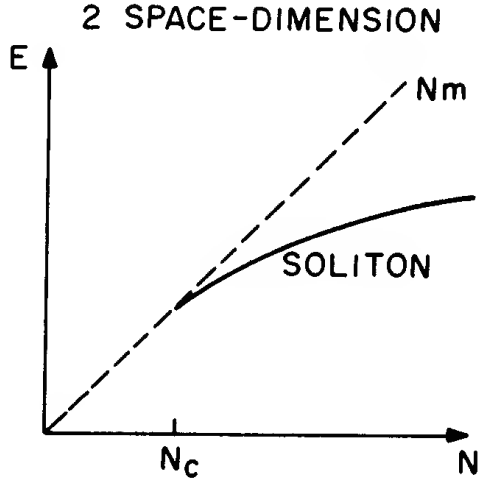


Fig. 7.14. The solid curve is E (soliton) vs. N for a two-space-dimensional nontopological soliton.

3. Three space-dimensions

As $\omega \to m-$ and when r is large, the asymptotic behavior of a nontopological soliton is still given by (7.59); the corresponding value of N is

$$N \to 2m \int |\phi|^2 d^3r \ .$$

Because of the exponential factor in (7.59), the d^3r integration gives a factor $\sim (m^2 - \omega^2)^{-3/2}$. Hence, as $\omega \to m-$,

$$N \sim O\left(\frac{1}{\sqrt{m^2 - \omega^2}}\right) \to \infty \ . \qquad (7.61)$$

It can also be shown* that when ω is sufficiently small, N has to

* R. Friedberg, T. D. Lee and A. Sirlin, Phys.Rev. D13, 2739 (1976), Nucl.Phys. B115, 1, 32 (1976).

increase with decreasing ω, and in that way we can approach the limit given by (7.58). This, together with

$$\frac{d}{dN} E(\text{soliton}) = \omega > 0 \tag{7.62}$$

produces the rather intriguing shape of the E(soliton) curve shown in Fig. 7.15. When $\omega = m-$, $N \to \infty$ but the soliton energy is $Nm+$.

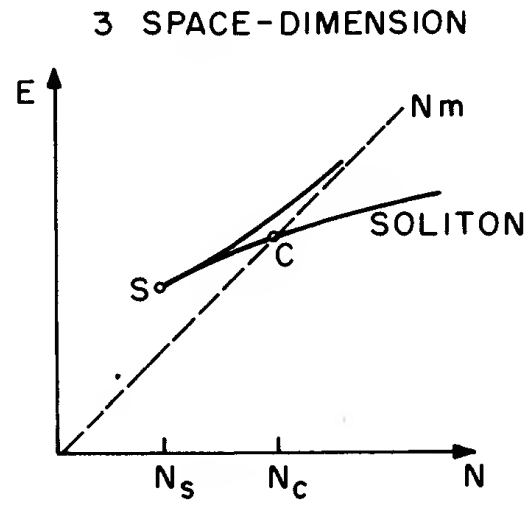


Fig. 7.15. The solid curve is E(soliton) vs. N for a three-space-dimensional non-topological soliton.

As ω decreases from $m-$, N also decreases until it reaches N_s; throughout this interval the soliton energy is $> Nm$. As ω decreases further, N then starts to increase; it crosses the line $E = Nm$ at $N = N_c$. If we decrease ω still further, then N keeps on increasing with E(soliton) always $< Nm$, until $N \to \infty$.

There now exists besides a critical point C also a spike S, with $N_c > N_s$. For $N < N_s$, there is no soliton solution. For $N_s < N < N_c$, the lowest-energy solution is the plane wave; for

$N > N_c$, the lowest-energy solution is always the soliton. In Fig. 7.15, along the lower branch of the soliton curve, for $N > N_c$, the soliton solution is absolutely stable, being the lowest-energy solution; for $N_c > N > N_s$, it can be shown that the soliton solution is stable against infinitesimal perturbations, even though it is not of the lowest energy. Along the upper branch, the soliton solution is always unstable. The numerical values of N_c and N_s depend on the parameters in the theory. It is not difficult to give examples in which N_c, and therefore also N_s, are < 1.

In a quantum theory N takes on only integer values $0, 1, 2, \cdots$. Thus, if $N_c < 1$, and if the classical results are good approximations of the quantum solutions, then the lowest-energy state of the system abruptly changes its character from the vacuum state ($N = 0$) to any $N \neq 0$ state. It is important to note that, when N is small, the classical soliton description of an N-body bound state is quite different from the usual description in terms of solutions of the Bethe-Salpeter equation (say, under the ladder approximation). For example, in the soliton description there is no sharp difference between the $N = 1$ and the $N = 2$ states; both are "blob-like". In the Bethe-Salpeter description, the $N = 1$ state is "point-like" while the $N = 2$ state is a "blob" formed by the wave function of the two "point-like" particles.

Remarks. In a nonlinear system, it is not difficult to conceive of situations in which the ground state can change its character quite suddenly with only a small variation in its parameters. The above three-space-dimensional case is one such example; when N varies from $< N_c$ to $> N_c$, the lowest-energy solution changes from

plane-wave to soliton. Similar examples can also be found in simple mechanical problems.

As an illustration, let us consider a single point particle moving in a central potential. The position vector of the particle is $\vec{r}$ and its conjugate momentum is $\vec{p}$. The Hamiltonian is assumed to be

$$\tfrac{1}{2}\vec{p}^{\,2} + \tfrac{1}{2} r^2[(1 - gr)^2 + \Delta^2] \tag{7.63}$$

where $r = |\vec{r}|$, and g and Δ are real parameters. The angular momentum $\vec{\ell} \equiv \vec{r} \times \vec{p}$ is conserved. At a fixed value of $\ell = |\vec{\ell}|$, (7.63) becomes

$$\tfrac{1}{2} p_r^{\,2} + V_\ell(r)$$

where p_r is the radial momentum and

$$V_\ell(r) = \tfrac{1}{2}\left(\frac{\ell}{r}\right)^2 + \tfrac{1}{2} r^2[(1 - gr)^2 + \Delta^2] \quad .$$

Here ℓ plays the same role as the conserved quantity N in our previous discussions. For ℓ and Δ both not too large and $g > 0$, $V_\ell(r)$ has two local minima, say at $r = r_1$ and r_2 with $r_1 < r_2$. As shown in Fig. 7.16(a), when $\ell = 0$ one sees that $r_1 = 0$. Hence,

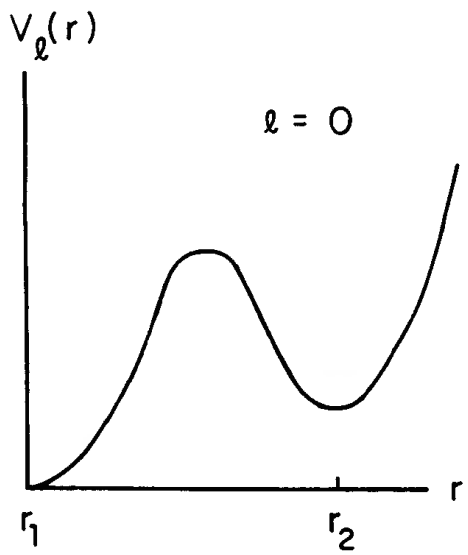


Fig. 7.16(a)

r_1 denotes the absolute minimum when ℓ is small. It is easy to show that there exists a critical value ℓ_c. For $\ell > \ell_c$, the absolute minimum of $V_\ell(r)$ changes from r_1 to r_2, as shown in Fig. 7.16(b). Now, in a quantum theory, ℓ takes on only integer values.

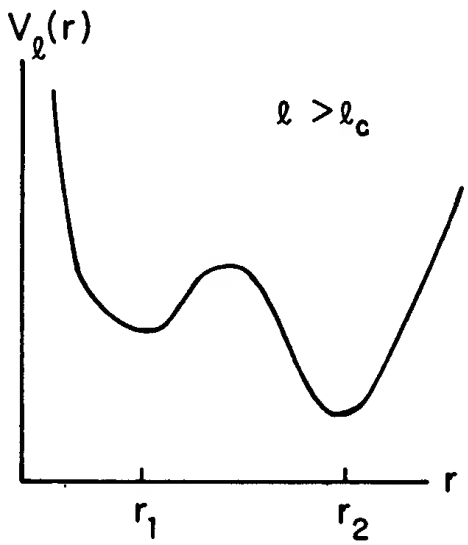


Fig. 7.16(b)

Fig. 7.16. $V_\ell(r)$ vs. r in a simple example for (a) $\ell = 0$ and (b) $\ell > \ell_c$.

Thus, if ℓ_c is < 1, the character of the $\ell = 0$ state can be drastically different from all $\ell \neq 0$ states. This is quite analogous to our field-theoretical problem, in which depending on the parameters, the vacuum state $(N = 0)$ may also be significantly different from all $N \neq 0$ states, provided that N_c is < 1.

Notice that in this example, if Δ is sufficiently small, one can always arrange to have $\ell_c < 1$, at any given g. A small g only means that the nonlinear solution is very far from the origin.

If N_c is small, then the soliton description of, say, the two-

body bound state is really quite different from the usual Bethe-Salpeter description (with, e.g., the ladder approximation).

7.6 Quantization

We now return to the important question of quantization, which was briefly commented upon in Section 7.2. There are many ways* to carry out the quantum expansion (7.12). Here we shall follow the canonical quantization procedure by using collective coordinates.**

1. Lagrangian, Hamiltonian and commutation relations

For simplicity, let us consider the one-space-dimensional non-topological soliton example discussed in Section 7.3. The Lagrangian density is given by (7.21)

$$\mathcal{L} = -\frac{\partial\phi^\dagger}{\partial x_\mu}\frac{\partial\phi}{\partial x_\mu} - \frac{1}{g^2}U(g^2\phi^\dagger\phi) \quad .$$

The corresponding Hamiltonian density is

$$\mathcal{H} = \Pi^\dagger\Pi + \frac{\partial\phi^\dagger}{\partial x}\frac{\partial\phi}{\partial x} + \frac{1}{g^2}U(g^2\phi^\dagger\phi) \quad , \tag{7.64}$$

where Π and $\Pi^\dagger$ are respectively the conjugate momenta of ϕ and $\phi^\dagger$, given by

$$\Pi = \frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \dot{\phi}^\dagger \quad \text{and} \quad \Pi^\dagger = \frac{\partial\mathcal{L}}{\partial\dot{\phi}^\dagger} = \dot{\phi} \quad . \tag{7.65}$$

The classical soliton solution can be written as

$$\phi_{cl} = \frac{1}{g}\,\sigma(x-\xi)\,e^{-i\theta} \tag{7.66}$$

where σ is given by (7.40), ξ is the integration constant and $\theta = \omega t$

---

* See the proceedings of the Conference on Extended Systems in Field Theory, Physics Reports 23C (1976) for detailed references.

** N. H. Christ and T. D. Lee, Phys.Rev. D12, 1606 (1975).

+ constant. The classical solution is degenerate under a constant variation in either ξ or θ; i.e., ϕ_{cl} remains a solution with the same energy under the variation

$$\phi_{cl} \rightarrow \phi_{cl} + \delta\phi \tag{7.67}$$

where

$$\delta\phi = -\frac{1}{g} e^{-i\theta} \left[\frac{d\sigma}{dx} \delta\xi + i\sigma \delta\theta \right] . \tag{7.68}$$

The variations $\xi \rightarrow \xi + \delta\xi$ represent a space translation and $\theta \rightarrow \theta + \delta\theta$ corresponds to a phase change in ϕ; their conjugate momenta are respectively the total momentum P and the particle number N of the system. The invariance under (7.67) - (7.68) implies that both P and N are conserved. Classically, ξ and θ commute with P and N; therefore, when the system has a definite momentum P and a definite particle number N, both ξ and θ can vary arbitrarily, resulting in the aforementioned degeneracy of the classical solution. As we shall see, quantum mechanically this degeneracy is lifted by treating ξ and θ as collective coordinates; their conjugate momenta become the derivative operators

$$P = \frac{1}{i} \frac{\partial}{\partial\xi} \qquad \text{and} \qquad N = \frac{1}{i} \frac{\partial}{\partial\theta} . \tag{7.69}$$

We shall first follow the standard procedure given in Chapter 2 for carrying out the quantization. The commutation relations are given by

$$[\Pi(x, t), \phi(x', t)] = [\Pi^\dagger(x, t), \phi^\dagger(x', t)] = -i\delta(x-x') \tag{7.70}$$

and all other equal-time commutators between ϕ, $\phi^\dagger$, Π and $\Pi^\dagger$ are 0. The dynamics of the system is then determined by requiring Heisenberg's equation

$$[H, O(t)] = -i \dot{O}(t)$$

to hold for any operator $O(t)$.

As in (2.48), we may decompose the complex field into its Hermitian components:

$$\phi = \frac{1}{\sqrt{2}} (\phi_1 + i\phi_2) \quad \text{and} \quad \phi^\dagger = \frac{1}{\sqrt{2}} (\phi_1 - i\phi_2) \tag{7.71}$$

where ϕ_1 and ϕ_2 are Hermitian fields. Their Fourier series can be written as

$$\phi_\alpha(x, t) = \sum_k \sqrt{\frac{2}{L}} \left[x_{\alpha,k}(t) \cos kx + y_{\alpha,k}(t) \sin kx \right] \tag{7.72}$$

where $\alpha = 1$ or 2, L is the volume (i.e., length) of the system and k is given by

$$k = \frac{2n\pi}{L} \quad \text{with} \quad n = 0, 1, 2, \cdots . \tag{7.73}$$

The Hermiticity of ϕ_α implies that $x_{\alpha,k}$ and $y_{\alpha,k}$ are also Hermitian; i.e.,

$$x_{\alpha,k} = x^\dagger_{\alpha,k} \quad \text{and} \quad y_{\alpha,k} = y^\dagger_{\alpha,k} . \tag{7.74}$$

The conjugate momenta Π and $\Pi^\dagger$ can be decomposed in a similar manner:

$$\Pi = \frac{1}{\sqrt{2}} (\Pi_1 - i\Pi_2) \quad \text{and} \quad \Pi^\dagger = \frac{1}{\sqrt{2}} (\Pi_1 + i\Pi_2) . \tag{7.75}$$

The Fourier components of Π_1 and Π_2 can be written in terms of the differential operators of $x_{\alpha,k}$ and $y_{\alpha,k}$. We have

$$\Pi_\alpha = -i \sum_k \sqrt{\frac{2}{L}} \left[\cos kx \frac{\partial}{\partial x_{\alpha,k}} + \sin kx \frac{\partial}{\partial y_{\alpha,k}} \right] . \tag{7.76}$$

Because of (7.72) and (7.76) we have

$$[\Pi_\alpha(x, t), \phi_\beta(x', t)] = -i\delta(x - x') \delta_{\alpha\beta} \tag{7.77}$$

and

$$[\phi_\alpha(x, t), \phi_\beta(x', t)] = [\pi_\alpha(x, t), \pi_\beta(x', t)] = 0$$

where α and β can be 1 or 2, which in turn lead to the original commutation relation (7.70).

In terms of these Fourier components, because

$$\int \dot{\phi}^\dagger \dot{\phi}\, dx = \sum_{\alpha,k} \tfrac{1}{2}(\dot{x}^2_{\alpha,k} + \dot{y}^2_{\alpha,k}) \ , \tag{7.78}$$

the Lagrangian density can be written as

$$\int \mathcal{L}\, dx = \sum_{\alpha,k} \tfrac{1}{2}(\dot{x}^2_{\alpha,k} + \dot{y}^2_{\alpha,k}) - V(x_{\alpha,k}, y_{\alpha,k}) \tag{7.79}$$

where

$$V = \int [\frac{\partial \phi^\dagger}{dx} \frac{\partial \phi}{\partial x} + \frac{1}{g^2} U(g^2 \phi^\dagger \phi)]\, dx \ . \tag{7.80}$$

Correspondingly, the Hamiltonian $H = \int \mathcal{H}\, dx$ is given by

$$H = -\tfrac{1}{2} \sum_{\alpha,k} (\frac{\partial^2}{\partial x^2_{\alpha,k}} + \frac{\partial^2}{\partial y^2_{\alpha,k}}) + V(x_{\alpha,k}, y_{\alpha,k}) \ . \tag{7.81}$$

While the Fourier series is a convenient expansion for analysing the plane wave, it is particularly ill-adapted for the soliton solution. We will therefore make a change of variables from the Fourier components to a new set.

2. Collective coordinates

The new set consists of two collective coordinates $\xi(t)$ and $\theta(t)$, introduced before, and a number of vibrational coordinates $q_{\pm,n}(t)$ with $n = 2, 3, \cdots$. If one wishes, one may label $q_{+,1}(t) = \xi(t)$ and $q_{-,1}(t) = \theta(t)$ so that the new set is $\{q_{\pm,n}(t)\}$. Let us choose two independent sets of complete, real and orthonormal c. number functions $\{\psi_{+,n}(x)\}$ and $\{\psi_{-,n}(x)\}$ with the conditions

$$\psi_{+,1}(x) = \text{constant times } \frac{d\sigma}{dx} \tag{7.82}$$

and

$$\psi_{-,1}(x) = \text{constant times } \sigma \ . \tag{7.83}$$

Hence for $n = 2, 3, \cdots$, we have

$$\int \frac{d\sigma}{dx} \psi_{+,n} dx = \int \sigma \psi_{-,n} dx = 0 \ . \tag{7.84}$$

Except for the above conditions, these two sets of c. number functions can be arbitrarily chosen. We now expand the quantum field operator $\phi(x, t)$ as follows:

$$\phi(x,t) = \{ \frac{1}{g} \sigma(x-\xi) + \sum_{n=2}^{\infty} \frac{1}{\sqrt{2}} [q_{+,n}(t) \psi_{+,n}(x-\xi) + i q_{-,n} \psi_{-,n}(x-\xi)] \} e^{-i\theta} . \tag{7.85}$$

These new coordinates $\xi(t)$, $\theta(t)$ and $q_{\pm,n}(t)$ are all Hermitian.

To understand the two conditions given in (7.84), let us expand ϕ around the soliton solution $\frac{1}{g}\sigma(x-\xi)e^{-i\theta}$ and regard $\delta\xi$, $\delta\theta$ and $q_{\pm,n}$ with $n \geqslant 2$ all as small. By using (7.68) and (7.85) we see that

$$\delta\phi = [-\frac{1}{g} (\frac{d\sigma}{dx} \delta\xi + i\sigma \delta\theta) + \sum_{n=2}^{\infty} 2^{-\frac{1}{2}} (q_{+,n} \psi_{+,n} + i q_{-,n} \psi_{-,n})] e^{-i\theta} . \tag{7.86}$$

Equation (7.84) insures that motions due to the variations in the collective variables ξ and θ will not be mistaken for the vibrational modes.

By equating (7.85) with the Fourier expansion (7.71) - (7.72), we obtain the coordinate transformation from

$$x_{\alpha,k}, \ y_{\alpha,k} \rightarrow q = \begin{pmatrix} \xi \\ \theta \\ q_{+,2} \\ q_{-,2} \\ \vdots \\ \vdots \end{pmatrix} \ . \tag{7.87}$$

This enables us to express the differential operators $\frac{\partial}{\partial x_{a,k}}$ and $\frac{\partial}{\partial y_{a,k}}$ as linear functions of $\frac{\partial}{\partial \xi}$, $\frac{\partial}{\partial \theta}$, $\frac{\partial}{\partial q_{+,2}}$, $\frac{\partial}{\partial q_{-,2}}$, $\cdots$. By substituting these linear relations into (7.81), we can obtain the Hamiltonian H in terms of these new variables. The easiest way to derive the explicit form is to first consider the classical problem. The time derivative of (7.85) gives $\dot{\phi}$ as a linear function of $\dot{\xi}$, $\dot{\theta}$, $\dot{q}_{\pm,2}$, $\cdots$; i.e.

$$\dot{\phi} = \sum_{\lambda,n} \frac{\partial \phi}{\partial q_{\lambda,n}} \dot{q}_{\lambda,n} \tag{7.88}$$

where

$$n = 1, 2, 3, \cdots, \qquad \lambda = +, - \ ,$$

and for notational convenience, we have set

$$q_{+,1} = \xi \qquad \text{and} \qquad q_{-,1} = \theta \ . \tag{7.89}$$

Thus, (7.78) can also be written as

$$\int \dot{\phi}^{\dagger} \dot{\phi} \, dx = \tfrac{1}{2} \tilde{\dot{q}} \, \mathcal{M} \dot{q} \tag{7.90}$$

where $\dot{q}$ is the time derivative of the column matrix in (7.87), and

$$\mathcal{M} = \mathcal{M}(q) \tag{7.91}$$

is a real, symmetric $\infty \times \infty$ matrix whose matrix elements are

$$\mathcal{M}_{\lambda,n;\lambda',n'}(q) = \int \left(\frac{\partial \phi^{\dagger}}{\partial q_{\lambda,n}} \frac{\partial \phi}{\partial q_{\lambda',n'}} + \frac{\partial \phi^{\dagger}}{\partial q_{\lambda',n'}} \frac{\partial \phi}{\partial q_{\lambda,n}} \right) dx \ . \tag{7.92}$$

By differentiating (7.90) with respect to $\dot{q}_{\lambda,n}$, we see that the conjugate momentum $p_{\lambda,n}$ of the generalized coordinate $q_{\lambda,n}$ is the $(\lambda, n)^{th}$ component of the column matrix

$$p = \mathcal{M} \dot{q} \ . \tag{7.93}$$

Classically, the Hamiltonian is

$$H_{cl} = \tfrac{1}{2} \tilde{p}_{cl} \, \mathcal{M}^{-1} \, p_{cl} + V(q) \tag{7.94}$$

where $\mathcal{m}^{-1}$ is the inverse of $\mathcal{m}$, and the function V(q) can be obtained by substituting (7.85) into (7.80). Quantum mechanically, because the transformation (7.87) is a point-transformation, the generalized Laplace operator in (7.81) satisfies (see Problem 7.3)

$$\sum_{a,k} \left(\frac{\partial^2}{\partial x_{a,k}^2} + \frac{\partial^2}{\partial y_{a,k}^2} \right) = \frac{1}{\mathcal{g}} \tilde{p}\, \mathcal{m}^{-1}\, \mathcal{g}\, p \qquad (7.95)$$

where the matrix element of the column matrix p is now the differential operator

$$p_{\lambda,n} = -i \frac{\partial}{\partial q_{\lambda,n}} \quad , \qquad (7.96)$$

$\mathcal{m}^{-1}$ remains the same inverse matrix in (7.94) and

$$\mathcal{g} = \sqrt{\det \mathcal{m}(q)} \quad . \qquad (7.97)$$

On account of (7.69) and (7.89), the first two components of p are the total momentum P and the particle number N of the system; i.e.

$$p_{+,1} = P = -i \frac{\partial}{\partial \xi} \quad \text{and} \quad p_{-,1} = N = -i \frac{\partial}{\partial \theta} \, . \qquad (7.98)$$

The quantum Hamiltonian (7.81) is equal to the Hermitian operator

$$H = \frac{1}{2\mathcal{g}} \tilde{p}\, \mathcal{m}^{-1}\, \mathcal{g}\, p + V(q) \quad . \qquad (7.99)$$

Because the classical soliton solution (7.66) is invariant under the variation (7.67)-(7.68) in the two collective coordinates ξ and θ, the functions V(q), $\mathcal{g}(q)$ and $\mathcal{m}(q)$ all do not depend on ξ and θ. Thus, P and N both commute with H :

$$[P, H] = [N, H] = 0 \quad , \qquad (7.100)$$

which means that they are both constants of motion. The eigenvalue of P is continuous, and is the total momentum of the system. Since θ is a cyclic variable, N has only discrete integer eigenvalues

$\cdots, -2, -1, 0, 1, 2, \cdots$, positive for particles and negative for anti-particles. The energy of the system is determined, as usual, by the Schrödinger equation

$$H \,|\,> = E \,|\,> \quad . \tag{7.101}$$

3. Perturbation expansion

To show how the eigenstate of H can be solved by the perturbation series in g, we assume that the vibrational coordinates $q_{\pm,2}$, $q_{\pm,3}$, $\cdots$ and their time derivatives are all $O(1)$. Thus, (7.85) gives

$$\dot{\phi} = -\frac{1}{g}\left(\frac{d\sigma}{dx}\,\dot{\xi} + i\sigma\dot{\theta}\right) e^{-i\theta} + O(1) \quad . \tag{7.102}$$

Hence (7.90) is $O(g^{-2})$, and it is given by

$$\int \dot{\phi}^{\dagger}\dot{\phi}\, dx = \frac{M}{2}\,\dot{\xi}^2 + \frac{I}{2}\,\dot{\theta}^2 + O(g^{-1}) \tag{7.103}$$

where

$$M = \frac{2}{g^2}\int \left(\frac{d\sigma}{dx}\right)^2 dx \quad \text{and} \quad I = \frac{2}{g^2}\int \sigma^2\, dx \quad . \tag{7.104}$$

By substituting (7.85) into (7.80), we find $V(q)$ is also $O(g^{-2})$:

$$V(q) = \frac{1}{g^2}\int \left[\left(\frac{d\sigma}{dx}\right)^2 + U(\sigma^2)\right] dx + O(g^{-1}) \quad . \tag{7.105}$$

To simplify our analysis, let us assume the soliton is at rest; i.e.,

$$P \,|\,> = 0 \quad . \tag{7.106}$$

By following the steps from (7.90)-(7.99) and by using (7.105)-(7.106), we see that

$$H = \frac{N^2}{2I} + \frac{1}{g^2}\int \left[\left(\frac{d\sigma}{dx}\right)^2 + U(\sigma^2)\right] dx + O(g^{-1}). \tag{7.107}$$

Let us define ω by setting the eigenvalue of N to be

$$N = I\omega \quad . \tag{7.108}$$

Because N is also the derivative of (7.103) with respect to $\dot{\theta}$, we see that $N = I\dot{\theta}$ and therefore ω is the same parameter that

characterizes the classical solution σ. The only restriction in the quantum solution is that now N must be an integer. Consequently, on account of (7.104) and (7.108), (7.107) can be written as

$$H = \frac{1}{g^2} \int \left[\left(\frac{d\sigma}{dx}\right)^2 + U(\sigma^2) + \omega^2 \sigma^2 \right] dx + O(g^{-1}) \quad . \tag{7.109}$$

By using the classical soliton solution (7.31), we find that its energy is of exactly the same form

$$E_{cl} = \frac{1}{g^2} \int \left[\left(\frac{d\sigma}{dx}\right)^2 + U(\sigma^2) + \omega^2 \sigma^2 \right] dx \quad ; \tag{7.110}$$

therefore, the quantum solution becomes

$$H \mid > = [E_{cl} + O(1)] \mid > \quad . \tag{7.111}$$

In (7.109), there seems to be an $O(g^{-1})$ correction term. However, because σ is the solution of the classical equation, it is not difficult to show that the $O(g^{-1})$ term is in fact zero, which leads to (7.111).

Thus, when g is small the existence of a classical soliton solution insures that of a quantum solution; furthermore, when $g \to 0$, the quantum soliton mass is given by the same classical curve $E_{cl}(N)$, except that N must be integers. As noted in (7.12) - (7.13), these properties have a general validity not restricted to the particular type of one-space-dimensional examples discussed in this section.

Remarks. In this introduction to field theory we have covered the basic quantization procedures and the general method of evaluating the S-matrix. The perturbation series around the plane-wave solutions will be useful to describe leptons and photons which have no strong interactions; as we shall see, there exist compelling reasons to regard all known hadrons as bound states of quarks. The quarks seem to be permanently confined in space and that is why free quarks are not

seen. At distances $\sim$ the hadron radius, the dynamics of these quark composites (i.e., hadrons) can be best described in terms of the soliton solution. However, at very small distances inside the hadron, because of the asymptotic freedom property of quantum chromodynamics, the usual perturbation expansion around plane-wave solutions again becomes approximately valid. These topics will be analysed later on.

Our discussions of particle physics in the subsequent chapters will consist of two main parts: symmetry and interactions. In symmetry, the analysis will be entirely phenomenological, and yet rigorous conclusions can be derived without any detailed knowledge of the dynamics. In the chapters dealing with interactions, specific assumptions will be made about the Hamiltonian, which enable us to make more detailed calculations.

We shall now go on to particle physics.

Problem 7.1. Let the Lagrangian density of a one-space-dimensional Hermitian field σ be

$$\mathcal{L} = -\tfrac{1}{2}\left(\frac{\partial \sigma}{\partial x_\mu}\right)^2 - V(\sigma)$$

where $x_\mu = (x, it)$. Show that the classical soliton solution for

(i) $V(\sigma) = \frac{1}{8}(1-\sigma^2)^2$ is $\sigma = \tanh\left[\frac{1}{2}(x-\xi)\right]$

and (ii) $V(\sigma) = 1 - \cos\sigma$ is $\sigma = 4\tan^{-1} e^{x-\xi}$, with ξ as the integration constant.

The field equation of (ii) is called the sine-Gordon equation.

Problem 7.2. Let the Lagrangian density of a one-space-dimensional complex field σ be

$$\mathcal{L} = -\frac{\partial \sigma^\dagger}{\partial x_\mu}\frac{\partial \sigma}{\partial x_\mu} - U(\sigma^\dagger \sigma) \quad .$$

Show that the classical soliton solution for

$$U = \frac{\sigma^\dagger \sigma}{1+\epsilon^2}\left[(1-\sigma^\dagger\sigma)^2 + \epsilon^2\right]$$

is

$$\sigma = \left(\frac{a}{1+\sqrt{1-a}\,\cosh y}\right)^{\frac{1}{2}} e^{-i\omega t} \qquad (7.112)$$

where $a = (1+\epsilon^2)(1-\omega^2)$ and $y = 2\sqrt{1-\omega^2}\,(x-\xi)$.

In both problems, for simplicity we set the scale so that $g = 1$ and $m = 1$.

Problem 7.3. In the transformation $x_i \to q_j = q_j(x_i)$ where the subscripts can vary from 1 to n, we define M to be the n x n matrix $M = (M_{ij}) \equiv \left(\frac{\partial x_k}{\partial q_i}\frac{\partial x_k}{\partial q_j}\right)$, $M^{-1} = (M^{-1}_{ij})$ its inverse and $|M| = \mathcal{J}^2$ its determinant. Show that

$$\text{(i)}\quad \frac{1}{|M|}\frac{\partial |M|}{\partial q_k} = M^{-1}_{ji}\frac{\partial M_{ij}}{\partial q_k} = -2\frac{\partial x_i}{\partial q_k}\frac{\partial}{\partial q_j}\left(\frac{\partial q_j}{\partial x_i}\right)$$

and

$$\text{(ii)}\quad \frac{\partial^2}{\partial x_i\,\partial x_i} = \frac{1}{\mathcal{J}}\frac{\partial}{\partial q_i}\left(M^{-1}_{ij}\,\mathcal{J}\,\frac{\partial}{\partial q_j}\right) . \qquad (7.113)$$

References

For quantization and especially topological solitons, there are some excellent review articles:

S. Coleman, in New Phenomena in Subnuclear Physics, Part A, ed. A. Zichichi (New York and London, Plenum Press, 1977), 297.

L. D. Faddeev and V. E. Korepin, Physics Reports C42, No. 1 (1978).

R. Jackiw, Revs. Mod. Phys. 49, 681 (1977).

Chapter 8

ORDER-OF-MAGNITUDE ESTIMATIONS

From a phenomenological point of view, there are four distinct interactions between particles, given in Table 1.

Table 1

Interaction	Phenomenological Coupling Constant	Quanta of the Mediating Field
Strong	~ 1	Color gluon
Electromagnetic	$\alpha \cong \frac{1}{137}$	Photon
Weak	$G m_p^2 \sim 10^{-5}$	Intermediate boson
Gravitational	$\mathcal{G} m_p^2 \sim 6 \times 10^{-39}$	Graviton

The four phenomenological classes of interactions. α is the fine structure constant, G is the Fermi constant for beta-decay, $\mathcal{G}$ is Newton's gravitational constant and m_p is the mass of the proton; all these constants are in the natural units. Among the quanta of the mediating fields, only the photon has been detected experimentally.

Throughout our discussions we will concentrate only on the strong, weak and electromagnetic interactions. Except for the quanta of the mediating fields, all particles can be classified into two types.

Those that have strong interactions are called hadrons, all others are leptons. For example, p, n, Λ, π are all hadrons; e, μ, τ and various neutrinos are leptons. (Since the mass of τ is about 2 GeV, the word 'lepton', which means light particle, is something of a misnomer.) The detailed properties of these particles are given in a table at the end of this book.

In this chapter we shall illustrate how to estimate the order of magnitude of physical quantities, such as various particle sizes and cross sections. These estimations will employ very few input parameters, listed below:

the fine structure constant	$a \cong \frac{1}{137}$,	
the Fermi constant	$G \cong 10^{-5}/m_p^2$,	
the electron mass	$m_e \cong 0.51\ \text{MeV} \cong (4\times10^{-11}\ \text{cm})^{-1}$	(8.1)
the proton mass	$m_p \cong 1800\ m_e$,	
and		
the pion mass	$m_\pi \cong \frac{1}{7} m_p$.	

The art of order-of-magnitude estimations is based on

(i) simple physical considerations,

and (ii) dimensional analysis,

as will be illustrated by the following examples.

8.1 Radius of the Hydrogen Atom

The hydrogen radius r is determined by the orbit of the electron. Therefore, the momentum of the electron is $p \sim \frac{1}{r}$ and its kinetic energy is $\sim \frac{p^2}{2m_e}$. The electrostatic energy should be $\sim -\frac{a}{r}$. The energy E can then be estimated to be

$$E \sim \frac{1}{2m_e}\left(\frac{1}{r}\right)^2 - \frac{\alpha}{r} \quad . \tag{8.2}$$

Its minimum is determined by

$$\frac{dE}{dr} = 0 \ ,$$

which leads to

$$r = \frac{1}{m_e \alpha} \quad . \tag{8.3}$$

By using the values of α and m_e in (8.1) we see that

$$r \cong 5 \times 10^{-9} \text{ cm} \tag{8.4}$$

which is the Bohr radius, now derived without solving any differential equations.

Remarks. In quantum electrodynamics there are three important lengths, differing from each other by powers of α :

$$\begin{array}{ll} \text{Bohr radius} & \dfrac{1}{m_e \alpha} \ , \\ \text{Compton wavelength of e} & \dfrac{1}{m_e} \ , \\ \text{classical radius of e} & \dfrac{\alpha}{m_e} \ . \end{array} \tag{8.5}$$

8.2 Hadron Size

According to Table 1, the strong interaction coupling constant is ~ 1 instead of α. Therefore a glance at (8.5) tells us, for the hadrons, there is only one length, which can be taken as the Compton wavelength of the particle. Since the pion is the lowest-mass hadron, its Compton wavelength is therefore the largest. Because of strong interactions, pion clouds must exist in other hadrons such as p,

n, Λ, $\cdots$. All these hadrons, including the pion, should be of a size $\sim \frac{1}{m_\pi}$ which according to (8.1) is $\sim 10^{-13}$ cm = 1 fermi.

Therefore, one expects the charge radius r_p of the proton to be of the same order. Experimentally, one finds

$$r_p \cong .81 \text{ fermi}, \tag{8.6}$$

consistent with the above simple estimation.

8.3 High-energy pp, πp and Kp Total Cross Sections

Because pp has strong interactions, we expect that at high energy the total cross section of a pp collision will be

$$\sigma_{pp} \sim \pi r_p^2 . \tag{8.7}$$

Since r_p is $\sim 10^{-13}$ cm, we estimate

$$\sigma_{pp} \sim 3 \times 10^{-26} \text{ cm}^2 \cong 30 \text{ mb} \tag{8.8}$$

where 1 mb = 10^{-3} b and 1 b = 1 barn = 10^{-24} cm^2.

As we shall see later in the discussion of the quark-parton model, a nucleon (proton or neutron) is made of three quarks while a meson (π or K) is made of two. Furthermore, at high energy we can treat these quarks as free particles, at least so far as total cross sections are concerned. Thus we expect

$$\frac{\sigma_{\pi p}}{\sigma_{pp}} \cong \frac{\sigma_{Kp}}{\sigma_{pp}} \cong \frac{2}{3}$$

which is independent of the charge of π and K. By using (8.8), we can estimate, for the total cross sections,

$$\begin{aligned} \sigma_{\pi^+ p} &\cong \sigma_{\pi^- p} \cong 20 \text{ mb} , \\ \sigma_{K^+ p} &\cong \sigma_{K^- p} \cong 20 \text{ mb} . \end{aligned} \tag{8.9}$$

Likewise, we expect

$$\sigma_{np} \cong \sigma_{pp} \cong \sigma_{\bar{p}p} \cong \sigma_{\bar{n}p} \quad . \tag{8.10}$$

The present high-energy experimental values are

$$\sigma_{pp} \cong \sigma_{np} \cong \sigma_{\bar{p}p} \cong \sigma_{\bar{n}p} \cong 45 \text{ mb} \quad ,$$

and

$$\sigma_{\pi^{\pm}p} \cong 25 \text{ mb}$$

$$\sigma_{K^{\pm}p} \cong 20 \text{ mb} \quad ,$$

all consistent with the above estimations.

8.4 $e^+ + e^- \to \mu^+ + \mu^-$

We shall estimate the total cross section for this reaction at high energy. Since $e^{\pm}$ and $\mu^{\pm}$ are leptons, the strongest interaction between them is the electromagnetic. The lowest order diagram is given in Fig. 8.1 where the wavy line represents the virtual photon.

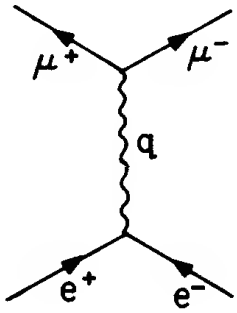


Fig. 8.1 Feynman diagram for $e^+ + e^- \to \mu^+ + \mu^-$.

Let σ be the total cross section. In Fig. 8.1 each vertex carries a factor proportional to the electric charge; the Feynman amplitude is therefore proportional to the fine structure constant α . So far as the

total cross section is concerned, the only Lorentz-invariant variable in this problem is the square of the 4-momentum q carried by the virtual photon. It is useful to visualize the process in the center-of-mass system (which is also the laboratory system if this is a standard colliding beam experiment), and to denote s as the square of the center-of-mass energy $E_{c.m.}$. In our case, we have

$$s \equiv E_{c.m.}^2 = -q^2 = q_0^2 - \vec{q}^{\,2} \ . \tag{8.11}$$

The total cross section must therefore be of the form

$$\sigma = \alpha^2 f(s, m_e, m_\mu) \tag{8.12}$$

where m_e and m_μ are respectively the masses of e and μ, and the function f is to be determined. When $E_{c.m.}$ is much greater than either lepton mass, we may set $m_e = m_\mu = 0$ as an approximation. Now, α is dimensionless, and in the natural units the dimensions of σ and s are

$$[\sigma] = [L^2] \qquad \text{and} \qquad [s] = [L^{-2}] \ , \tag{8.13}$$

where L denotes length. Thus, from dimensional analysis the function f in (8.12) for large s must be proportional to s^{-1}; i.e.

$$\sigma \sim \frac{\alpha^2}{s} \ . \tag{8.14}$$

The Feynman diagram in Fig. 8.1 can be readily evaluated, and the complete answer for σ in the high-energy limit is

$$\sigma = \frac{4}{3}\,\pi\,\frac{\alpha^2}{s} \ . \tag{8.15}$$

[See Problem 6.1.] The point is that without any computation it is possible to give an estimation of the cross section, albeit without the factor $\frac{4}{3}\,\pi$.

From (8.14), one can readily estimate, e.g., that when $s \cong (1\ \text{GeV})^2$

$$\sigma(e^+e^- \to \mu^+\mu^-) \sim 4 \times 10^{-32}\ \text{cm}^2 = .04\ \mu\text{b} \quad , \qquad (8.16)$$

where $1\ \mu\text{b} = 1$ microbarn $= 10^{-6}$ b .

8.5 $\nu + N \to \cdots$

Let $\sigma(\nu N)$ be the total cross section of this reaction summed over all final states. The initial ν can be either the μ-neutrino ν_μ or the e-neutrino ν_e, and N is the nucleon. As before, s denotes the square of the center-of-mass energy. Since this is a weak process, its amplitude should be proportional to the Fermi constant G. Hence, $\sigma(\nu N)$ must be of the form

$$\sigma(\nu N) = G^2 f(s, m_N) \qquad (8.17)$$

where m_N is the nucleon mass and the function f is to be determined. For $s \gg m_N^2$, we can, as in the preceding example, set $m_N = 0$ as an approximation. Since the dimension of G is

$$[G] = [L^2] \quad , \qquad (8.18)$$

by using (8.13) we see that in (8.17) the function f must be proportional to s ; therefore,

$$\sigma(\nu N) \sim G^2 s \quad . \qquad (8.19)$$

The laboratory system in a typical high-energy neutrino experiment is one in which the initial nucleon is at rest. Let E_ν and $\vec{p}_\nu$ be respectively the energy and the momentum of the neutrino in the laboratory system. We have

$$s = E_{c.m.}^2 = (E_\nu + m_N)^2 - \vec{p}_\nu^{\,2} = m_N(2E_\nu + m_N) \cong 2m_N E_\nu \, . \qquad (8.20)$$

Thus, (8.19) can also be written in terms of E_ν :

$$\sigma(\nu N) \sim G^2 m_N E_\nu \quad . \tag{8.21}$$

By using (8.1), we find at high energy

$$\sigma(\nu N) \sim 4 \times 10^{-38} \left(\frac{E_\nu}{m_N}\right) \text{cm}^2 \quad . \tag{8.22}$$

The experimental result is

$$\sigma(\nu N) \cong 0.6 \times 10^{-38} \left(\frac{E_\nu}{m_N}\right) \text{cm}^2 \quad . \tag{8.23}$$

Again, we can estimate the order of magnitude of this reaction without any computation. [Notice the huge difference between (8.8), (8.16) and (8.22) at a comparable energy.]

8.6 Compton Scattering

The lowest-order Feynman diagrams for Compton scattering

$$\gamma + e^{\pm} \rightarrow \gamma + e^{\pm} \tag{8.24}$$

are given in Fig. 8.2. By following the same reasoning which led to (8.12), we expect the total cross section σ_{Comp} of this reaction to

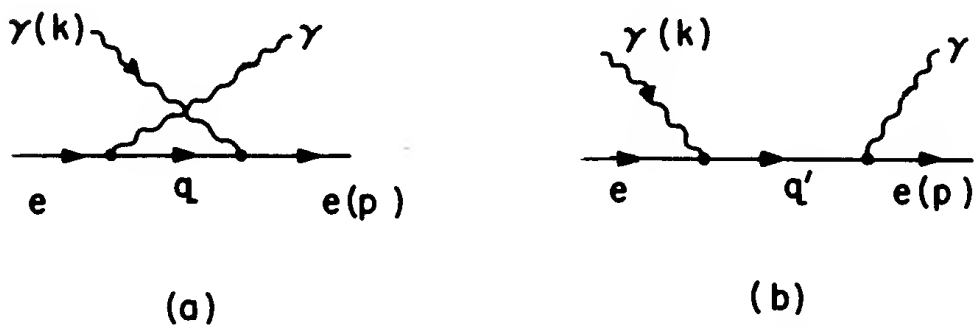


Fig. 8.2. Lowest-order diagrams for Compton scattering in which p and k denote the 4-momenta of the final e and the initial γ.

be of the form

$$\sigma_{Comp} = \alpha^2 f(s, m_e) \tag{8.25}$$

where, as before, s is the center-of-mass energy squared, and the function f is to be determined.

At the nonrelativistic limit, $s \to m_e^2$ and therefore f is a function which depends only on m_e. From dimensional considerations, we expect

$$\sigma_{Comp} \sim \frac{\alpha^2}{m_e^2} \qquad \text{N. R.} \tag{8.26}$$

where N.R. denotes the nonrelativistic limit. When s is $\gg m_e^2$, we can neglect m_e in (8.25), and that leads to, through dimensional analysis,

$$\sigma_{Comp} \sim \frac{\alpha^2}{s} \qquad \text{E. R.} \tag{8.27}$$

where E.R. denotes the extreme relativistic limit.

An accurate calculation of the Feynman diagrams in Fig. 8.2 gives *

$$\sigma \cong \begin{cases} \frac{8}{3}\pi \left(\frac{\alpha}{m_e}\right)^2 & \text{N. R.} \\ 2\pi \frac{\alpha^2}{s} \ln \frac{s}{m_e^2} & \text{E. R.} \end{cases} . \tag{8.28}$$

Our estimation (8.26) differs from the accurate nonrelativistic formula (called the Thomson limit) only by a factor $\frac{8\pi}{3}$. However, in the extreme relativistic case, the estimation (8.27) misses an s-dependent factor $2\pi \ln \frac{s}{m_e^2}$. While this is a slowly varying function of s, the existence of such log terms has a general underlying reason, as we shall explain.

---

* See W. Heitler, The Quantum Theory of Radiation (Oxford, Oxford University Press, 1944).

8.7 Mass Singularity and High-energy Behavior

We first give the technical reason for the $\ln \frac{s}{m_e^2}$ factor. Let us consider diagram (a) in Fig. 8.2. The virtual electron carries a 4-momentum

$$q = p - k \ , \tag{8.29}$$

where p is the final 4-momentum of e and k the initial 4-momentum of γ. The components of p and k may be written as $p = (\vec{p}, ip_0)$ and $k = (\vec{k}, ik_0)$. Since the final e and the initial γ are both on the mass shell, we have

$$p^2 + m_e^2 = 0 \quad \text{and} \quad k^2 = 0 \ ,$$

i.e.

$$p_0 = \sqrt{\vec{p}^{\,2} + m_e^2} \quad \text{and} \quad k_0 = |\vec{k}| \ . \tag{8.30}$$

For definiteness, we consider the laboratory frame. In the E.R. limit, p_0 is $\gg m_e$, and therefore

$$p_0 \cong |\vec{p}| + \frac{m_e^2}{2p_0} \ . \tag{8.31}$$

In diagram (a) the electron propagator carries a denominator

$$q^2 + m_e^2 = (p - k)^2 + m_e^2 \tag{8.32}$$

which, because of (8.30), is equal to

$$-2p \cdot k = -2\vec{p} \cdot \vec{k} + 2p_0 k_0 \ .$$

In the E.R. limit and for the nearly forward scattering case, we have

$$\begin{aligned} q^2 + m_e^2 &\cong m_e^2 \frac{k_0}{p_0} + 2|\vec{p}|\, k_0(1 - \cos\theta) \\ &\cong \frac{k_0}{p_0}(m_e^2 + p_0^2\, \theta^2) \end{aligned} \tag{8.33}$$

where $\theta \ll 1$ is the angle between $\vec{p}$ and $\vec{k}$. Diagram (a) gives a contribution σ_a to the cross section that is inversely proportional to

$(q^2 + m_e^2)^2$. For the region under consideration we expect the deviation of σ_a from the simple estimate (8.27) to be

$$\frac{\sigma_a}{(\alpha^2/s)} \sim \int \frac{\theta^2 2\pi \sin\theta \, d\theta}{(m_e^2 + p_0^2 \theta^2)^2} \cdot p_0^4$$

$$\sim 2\pi \ln\left(\frac{p_0}{m_e}\right)^2 \sim 2\pi \ln\left(\frac{s}{m_e^2}\right) \tag{8.34}$$

where $2\pi \sin\theta \, d\theta$ is the solid angle, p_0^4 is there to make the whole expression dimensionless, and the θ^2 factor in the numerator is due to γ_5 invariance, as will be explained below. The new estimation (8.34) is in good agreement with the exact limit given by the Klein-Nishina formula * given in (8.28).

The γ_5-transformation

$$\psi \rightarrow \gamma_5 \psi \ , \tag{8.35}$$

has been discussed in Section 3.8 in connection with the two component theory. For the electron field ψ, since $m_e \neq 0$, the mass term violates the γ_5 invariance. However, the electromagnetic interaction is invariant under the γ_5 transformation, as can be readily verified by substituting (8.35) into (6.7). In the E.R. limit, m_e can be neglected; therefore, by following the discussions given in Section 3.8 one sees that the helicity of the electron is unchanged through its electromagnetic interaction. Since the helicity of the photon is either $+1$ or -1, a helicity-conserving electron cannot emit or absorb a photon in the exact forward direction, as can be easily seen through angular-momentum conservation along the direction of motion. Thus, for zero m_e when the angle θ between $\vec{p}$ and $\vec{k}$ is 0, the

---

* O. Klein and Y. Nishina, Zeit.f.Phys. 52, 853 (1929).

Feynman amplitude must also be 0. Consequently, the matrix element for diagram (a) in Fig. 8.2 carries a factor proportional to θ, and that explains the θ^2 term in (8.34). [For $m_e \neq 0$ but $\theta = 0$, the amplitude is $\propto m_e$; terms proportional to m_e do not lead to the mass singularity.]

From (8.34) one observes that in the E.R. limit the cross section has a logarithmic singularity (called mass singularity) when the electron mass $m_e \to 0$. The origin of the mass singularity is associated with the fact that (8.33) becomes 0 when $m_e = \theta = 0$; i.e., the relevant propagator becomes infinite, which in turn means the virtual particle is approaching its mass shell. There is a simple, but general, reason for this: If we have a number of zero-mass particles moving in the same direction, their total energy $E = \sum E_a$ and the magnitude of their total 3-momentum $\vec{p} = \sum \vec{p}_a$ become equal

$$\sum_a E_a = \left| \sum \vec{p}_a \right| \tag{8.36}$$

where the subscript a denotes the a^{th} particle. Thus, when $m_e \to 0$, conservation of energy follows from the conservation of momentum whenever the momenta of e and γ are all parallel; i.e., in

$$e \rightleftarrows e + \gamma$$

all particles can be on-mass-shell in this special limit.

A general theorem can be established which states that, for any process

$$i \to f \ , \tag{8.37}$$

although the Feynman diagrams have mass singularities, the square of the S-matrix element

$$\sum_{\{i\},\{f\}} |\langle f | S | i \rangle|^2 \ , \tag{8.38}$$

summed over the sets $\{i\}$ and $\{f\}$ which respectively contain all states that are degenerate (within an energy width $\epsilon \neq 0$, which can be arbitrarily chosen) with i and f, does not. Expression (8.38) depends on the width ϵ but has a finite limit when $m_e \to 0$ to every order in the perturbation expansion. This theorem is valid not only in quantum electrodynamics, but also in other field theories such as quantum chromodynamics. The proof is quite elementary, and is given in Chapter 23.

8.8 e^+e^- Pair Production by High-energy Photons

Let us consider the reaction

$$\gamma + Z \to e^+ + e^- + Z \tag{8.39}$$

where Z denotes a heavy nucleus of charge Z (in units of positron charge). The lowest-order diagrams are given in Fig. 8.3 in which the wavy lines represent real or virtual photons. In order to establish the cross section of (8.39) we first consider the reaction of pair production by two photons

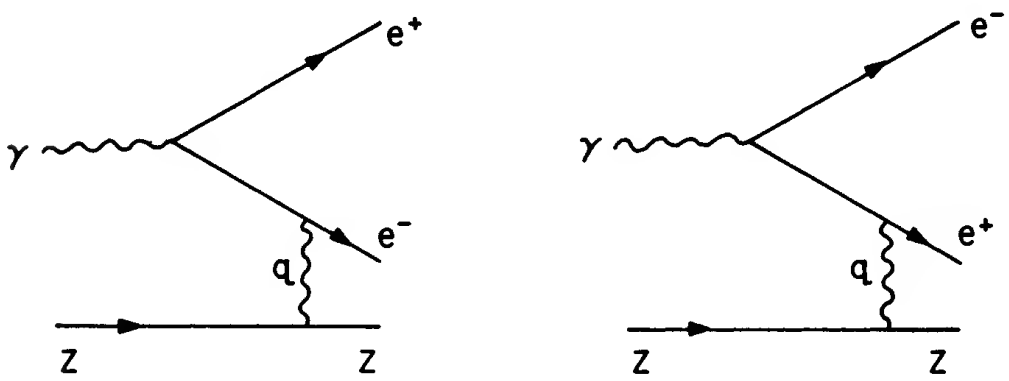


Fig. 8.3. Lowest-order Feynman diagrams for (8.39).

$$\gamma + \gamma \rightarrow e^+ + e^- \tag{8.40}$$

whose diagram is given in Fig. 8.4. By following the same argument

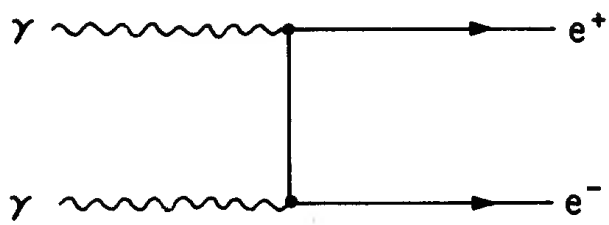


Fig. 8.4. Lowest-order diagram for $\gamma + \gamma \rightarrow e^+ + e^-$.

which leads to (8.27), we may estimate at high energy

$$\sigma(\gamma\gamma \rightarrow e^+e^-) \sim \frac{\alpha^2}{s} \ . \tag{8.41}$$

Next we compare the difference between reactions (8.39) and (8.40). For simplicity the nucleus is assumed to be extremely heavy, at rest in the laboratory system. The electrostatic potential generated by the nucleus at distance r is

$$\frac{Ze}{r} \ . \tag{8.42}$$

Thus, in the laboratory frame, the 4-momentum carried by the virtual photon in Fig. 8.3 is $q = (\vec{q}, 0)$; i.e., the fourth component of q is 0. The distribution of $\vec{q}$ is given by the Fourier transformation of (8.42), and therefore it is proportional to

$$\frac{1}{\vec{q}^2} \ . \tag{8.43}$$

The process $\gamma + Z \rightarrow e^+ + e^- + Z$ can then be viewed as a particular case of $\gamma + \gamma \rightarrow e^+ + e^-$ in which one of the γ's is virtual with a momentum distribution given by (8.43) in the laboratory frame.

Accordingly, in reaction (8.39) the 3-momenta $\vec{p}_+$ and $\vec{p}_-$ of e^+ and e^- are independent; therefore, the final-state phase space differs from that of (8.40) by an additional factor

$$\frac{d^3\vec{p}}{p_0} \sim \vec{p}^2 \sim s \tag{8.44}$$

where $p_0 = (\vec{p}^2 + m_e^2)^{\frac{1}{2}}$ and $\vec{p}$ can be either $\vec{p}_+$ or $\vec{p}_-$. By multiplying (8.44) and (8.41) we see that the s-factors cancel, and therefore at high energy the cross section σ_{pair} for reaction (8.39) is approximately independent of s. In Fig. 8.3, each Feynman diagram has three vertices, one is proportional to Ze and the other two to e. Thus, σ_{pair} is of the form

$$\sigma_{pair} \sim Z^2\alpha^3 f(m_e)$$

where the function f can be determined through dimensional analysis. We then obtain the estimation

$$\sigma_{pair} \sim \frac{Z^2\alpha^3}{m_e^2} \quad . \tag{8.45}$$

So far, we have only made the crudest estimation, ignoring completely the possibility of mass singularity. By following arguments similar to that given between (8.29) and (8.34), we can derive a better high-energy estimation

$$\sigma_{pair} \sim \frac{Z^2\alpha^3}{m_e^2} \ln \frac{E_\gamma}{m_e} \tag{8.46}$$

where E_γ is the laboratory energy of the initial γ in reaction (8.39). The exact limit given by the Bethe-Heitler formula* is

$$\sigma_{pair} = \frac{28}{9} \frac{Z^2\alpha^3}{m_e^2} \ln \frac{E_\gamma}{m_e} \quad , \tag{8.47}$$

---

* H. Bethe and W. Heitler, Proc.Roy.Soc. 146, 83 (1934). See also Heitler, The Quantum Theory of Radiation, loc. cit.

consistent with the above estimation.

Equation (8.47) is valid for a point nucleus without screening. If one has complete screening, then the log factor in (8.47) should be replaced by a constant $\cong \ln(183\ Z^{-\frac{1}{3}})$. Unlike σ_{Comp}, σ_{pair} does not decrease with increasing energy. This has an important experimental consequence; it means that in matter a high-energy photon has a finite, nonzero, almost energy-independent mean free path, which is, e.g., about 460 meters in air (1 atm. pressure and 0° C), 13 cm in aluminum and 7 mm in lead.

II A. PARTICLE PHYSICS: SYMMETRY

Chapter 9

GENERAL DISCUSSION

Since the beginning of physics, symmetry considerations have provided us an extremely powerful and useful tool in our effort to understand nature. Gradually they have become the backbone of our theoretical formulation of physical laws. In this chapter we shall review these symmetry operations and examine their foundation. Such an examination is useful, especially in view of the various asymmetries that have been discovered during the past quarter century.

There are four main groups of symmetries, or broken symmetries, that are found to be of importance in physics:

1. Permutation symmetry: Bose-Einstein and Fermi-Dirac statistics.
2. Continuous space-time symmetries, such as translation, rotation, acceleration, etc.
3. Discrete symmetries, such as space inversion, time reversal, particle-antiparticle conjugation, etc.
4. Unitary symmetries, which include

   U_1 - symmetries such as those related to conservation of charge, baryon number, lepton number, etc.,

   SU_2 (isospin) - symmetry,

   SU_3 (color) - symmetry,

and SU_n (flavor) - symmetry.

Among these, the first two groups, together with some of the U_1-symmetries and perhaps the SU_3 (color)-symmetry in the last group, are believed to be exact. All the rest seem to be broken.

9.1 Non-observables, Symmetry Transformations and Conservation Laws

The root of all symmetry principles lies in the assumption that it is impossible to observe certain basic quantities; these will be called "non-observables". Let us illustrate the relation between non-observables, symmetry transformations and conservation laws by a simple example. Consider the interaction energy V between two particles at positions $\vec{r}_1$ and $\vec{r}_2$. The assumption that the absolute position is a non-observable means that we can arbitrarily choose the origin O from which these position vectors are drawn; the interaction energy should be independent of O. In other words, V is invariant under an arbitrary space translation, changing O to O';

$$\vec{r}_1 \to \vec{r}_1 + \vec{d} \quad \text{and} \quad \vec{r}_2 \to \vec{r}_2 + \vec{d} \ , \tag{9.1}$$

as shown in Fig. 9.1. Consequently, V is a function only of the

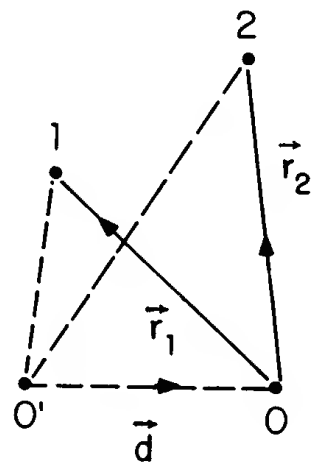


Fig. 9.1. The interaction energy V between particles 1 and 2 is invariant under a change of origin $O \to O'$.

relative distance $\vec{r}_1 - \vec{r}_2$,

$$V = V(\vec{r}_1 - \vec{r}_2) \quad . \tag{9.2}$$

From this, we deduce that the total momentum of this system of two particles must be conserved, since its rate of change is equal to the force

$$-(\vec{\nabla}_1 + \vec{\nabla}_2)\, V$$

which, on account of (9.2), is zero.

This example illustrates the interdependence among three aspects of a symmetry principle: the physical assumption of a non-observable, the implied invariance under the connected mathematical transformation and the physical consequences of a conservation law or selection rule. In an entirely similar way, we may assume the absolute time to be a non-observable. The physical law must then be invariant under a time-translation

$$t \rightarrow t + \tau$$

which results in the conservation of energy. By assuming the absolute spatial direction to be a non-observable, we derive rotation invariance and obtain the conservation law of angular momentum. By assuming that absolute (uniform) velocity is not an observable, we derive the symmetry requirement of Lorentz invariance, and with it the conservation laws connected with the six generators of the Lorentz group. Similarly, the foundation of general relativity rests on the assumption that it is impossible to distinguish the difference between an acceleration and a suitably arranged gravitational field.

The following table summarizes these three fundamental aspects for some of the symmetry principles used in physics.

Nan-abservables	Symmetry Transfarmations	Conservation Laws or Selection Rules
difference between identical particles	permutation	B.-E. or F.-D. statistics
absalute spatial pasition	space translatian $\vec{r} \to \vec{r} + \vec{\Delta}$	mamentum
absolute time	time translation $t \to t + \tau$	energy
absalute spatial directian	ratatian $\hat{r} \to \hat{r}'$	angular mamentum
absolute velocity	Lorentz transformatian	generators of the Larentz graup
absolute right (ar absalute left)	$\vec{r} \to -\vec{r}$	parity
absolute sign of electric charge	$e \to -e$ (or $\psi \to e^{i\Phi}\psi^{\dagger}$)	charge conjugatian (or particle anti-particle canjugation)
relative phase between states of different charge Q	$\psi \to e^{iQ\theta}\psi$	charge
relative phase between states af different baryan number N	$\psi \to e^{iN\theta}\psi$	baryon number
relative phase between states af different leptan number L	$\psi \to e^{iL\theta}\psi$	leptan number
difference between different coherent mixture af p and n states	$\begin{pmatrix} p \\ n \end{pmatrix} \to u\begin{pmatrix} p \\ n \end{pmatrix}$	isospin

Table 9.1.

9.2 Asymmetries and Observables

Since the validity of all symmetry principles rests on the theoretical hypothesis of non-observables, the violation of symmetry arises whenever what was thought to be a non-observable turns out to be actually an observable. In a sense, the discovery of "violations" is not that surprising. It is not difficult to imagine that some of the "non-observables" may indeed be fundamental, but some may simply be due to the limitations of our present ability to measure things. As we improve our experimental techniques, our domain of observation also enlarges. It should not be completely unexpected that we may succeed in detecting one of those supposed "non-observables" at some time and therein lies the root of symmetry breaking.

The notable examples of such discoveries are the asymmetry of physical laws under the right-left mirror transformation, the particle-antiparticle conjugation and the change in the direction of time flow, past to future and future to past. It turns out that all these supposed non-observables can actually be observed. Let me illustrate the relation between "asymmetries and observables" by first considering the example of right-left asymmetry, commonly known as parity nonconservation.

Of course it is well known that even in daily life, right and left are distinct from each other. Our hearts, for example, are usually not on the right side. The word "right" also means correct, while the word "sinister" in its Latin root means left. In English, one says "right-left", but in Chinese 左右 : 左 (left) usually precedes 右 (right). However, such asymmetry in daily life is attributed to either the accidental asymmetry of our environment or initial conditions. Before the discovery of parity nonconservation in 1957, it was taken for granted

that the laws of nature are symmetric under a right-left transformation.

One may wonder why, in spite of the clear difference between right and left in daily life, before 1956 practically all physicists could believe in right-left symmetry in physical laws.

Let us consider two cars which are made exactly alike, except that one is the mirror image of the other, as shown in Fig. 9.2. Car a

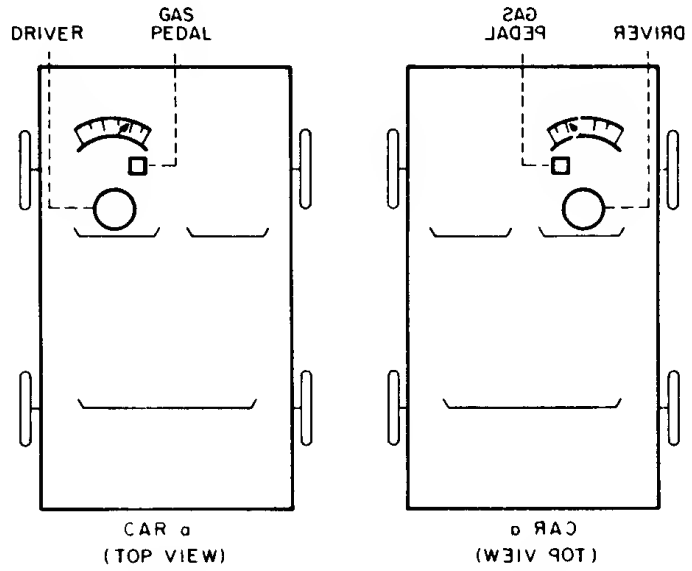


Fig. 9.2. Top view of two cars that are mirror images of each other.

has a driver on the left front seat and the gas pedal to his right, while b has the driver on the right front seat with the gas pedal on his left. Both cars are filled with the same amount of gasoline. (For the sake of discussion, we may ignore the fact that the gasoline molecule is not exactly mirror-symmetric.) Now, suppose the driver of Car a starts the car by turning the ignition key clockwise and stepping on the gas pedal with his right foot, causing the car to move forward at a certain

speed, say 20 mph. The other does exactly the same thing, except that he interchanges right with left; i.e., he turns the ignition key counterclockwise and steps on the gas pedal with his left foot, but keeps the pedal at the same degree of inclination. What will the motion of Car b be? The reader is encouraged to make a guess.

Probably your common sense will say that both cars should move forward at exactly the same speed. If so, you are just like the pre-1956 physicists. It would seem reasonable that two arrangements, identical except that one is the mirror image of the other, should behave in exactly the same way in all respects (except of course for the original right-left difference). This is precisely what was discovered to be untrue. The possibility of right-left asymmetry in natural laws was first suggested theoretically* in 1956 in connection with the θ-τ puzzle (see Chapter 15). It was discovered experimentally within a few months in β-decay** and in π- and μ-decays***. In principle, in the above example of two cars one may install, say, Co^{60} β-decay as part of the ignition mechanism. It will then be possible to construct two cars that are exact mirror images of each other, but may nevertheless move in a completely different way; Car a may move forward at a certain speed while Car b may move at a totally different speed, or may even move backwards. That is the essence of right-left asymmetry, or parity-nonconservation.

---

* T. D. Lee and C. N. Yang, Phys.Rev. 104, 254 (1956).

** C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson, Phys.Rev. 105, 1413 (1957).

*** R. L. Garwin, L. M. Lederman and M. Weinrich, Phys.Rev. 105, 1415 (1957); V. L. Telegdi and A. M. Friedman, Phys.Rev. 105, 1681 (1957).

As we shall discuss, the discoveries made in 1957 established not only right-left asymmetry, but also the asymmetry between the positive and negative signs of electric charge. In the standard nomenclature, right-left asymmetry is referred to as P violation, or parity nonconservation. The asymmetry between opposite signs of electric charge is called C violation, or charge conjugation violation, or sometimes particle-antiparticle asymmetry.

At the same time as the possibility of P and C violations was suggested, questions of possible asymmetries under time reversal and under CP were also raised.* Actual experimental confirmation did not come until quite a few years later.**

Since non-observables imply symmetry, these discoveries of asymmetry must imply observables. One may ask, what are the observables that have been discovered in connection with these symmetry-breaking phenomena? We recall that in our daily lives the sign of electric charge is merely a convention. The electron is considered to be negatively charged because we happened to assign the proton a positive charge, and the converse is also true. But now, with the discovery of asymmetry, is it possible to give an absolute definition? Can we find some absolute difference between the positive and negative signs of electricity, or between left and right?

As an illustration, let us consider the example of two imaginary, advanced civilizations A and B (see Fig. 9.3). These two civilizations are assumed to be completely separate from each other; nevertheless they manage to communicate, but only through neutral

---

* T. D. Lee, R. Oehme and C. N. Yang, Phys.Rev. 106, 340 (1957).

** J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, Phys. Rev.Lett. 13, 138 (1964).

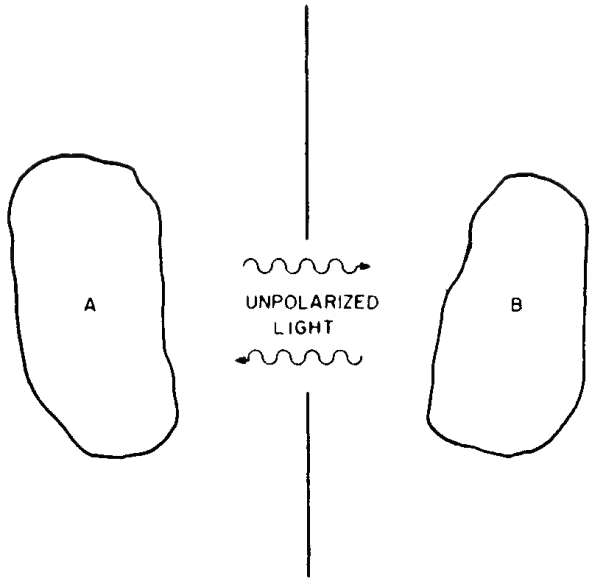


Fig. 9.3. Two imaginary civilizations communicate with each other through neutral unpolarized messages.

unpolarized messages, such as unpolarized light. After years of such communication these two civilizations may decide to increase their contact. Being very advanced, they realize that they must first agree on (i) the sign of electric charge,
and (ii) the definition of a righthand screw.
The first is important in order to establish whether the proton in civilization A corresponds to the proton or the antiproton in civilization B. If the protons in A are the same as those in B, then a closer contact is possible. Otherwise it might lead to annihilation. The definition of a righthand screw is important if these two civilizations decide to have even closer contact, such as trading machinery. The academic problem that concerns us is whether it is possible to transmit both pieces of information by using only neutral and unpolarized messages. Without these discoveries of P, C and CP asymmetries, this would

not be feasible. Now, assuming that these two civilizations are as advanced as ours, such an agreement can in principle be achieved.

First, both civilizations should establish high-energy physics laboratories which can produce the long-lived neutral kaon K_L^0. By analyzing the semileptonic three-body decay modes of K_L^0 under a magnetic field, they can easily separate the decay $K_L^0 \to e^- + \pi^+ + \bar{\nu}$ from $K_L^0 \to e^+ + \pi^- + \nu$. They would discover that although the parent particle K_L^0 is neutral and spherically symmetric, nevertheless these two decay rates are different

$$\frac{\text{rate}\ (K_L^0 \to e^+ + \pi^- + \nu)}{\text{rate}\ (K_L^0 \to e^- + \pi^+ + \bar{\nu})} = 1.00648 \pm 0.00035. \qquad (9.3)$$

This is indeed remarkable since it means that by rate counting one can differentiate e^- from e^+. Thus, there is an absolute difference between the opposite signs of electric charge. Now, each civilization only needs to examine the faster decay mode in (9.3), and compare the charge of the final e with that of its "proton." If both civilizations have the same relative sign, then it means that they are made of the same matter.

Next, we come to the second task: the definition of a right-hand screw. This can be done by measuring the spin and momentum direction of the neutrino or antineutrino in π-decay: $\pi^+ \to e^+ + \nu$ and $\pi^- \to e^- + \bar{\nu}$. Although $\pi^\pm$ is spherically symmetric, in its decay every ν defines a lefthand screw, while every $\bar{\nu}$ a righthand screw; i.e., the helicity of ν is always $-\frac{1}{2}$ and that of $\bar{\nu}$, $+\frac{1}{2}$, as shown in Fig. 9.4. The neutrino and antineutrino can therefore be described* by the two-component theory, discussed in Section 3.8.

---

* T. D. Lee and C. N. Yang, Phys. Rev. 105, 1671 (1957); L. Landau, Nucl. Phys. 3, 127 (1957); A. Salam, Nuovo Cimento 5, 299 (1957).

Consequently, we see that through neutral and unpolarized messages these two civilizations can indeed give an absolute meaning to + and - signs of charge as well as to right and left.

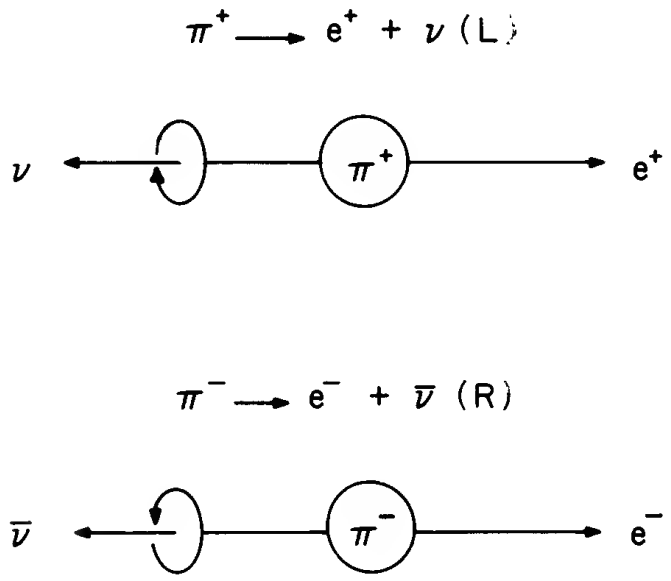


Fig. 9.4. The neutrino is a perfect lefthand screw, while its antiparticle is a perfect righthand screw.

We note that in the decay $\pi^{\pm} \rightarrow e^{\pm} + \nu$ (or $\bar{\nu}$), both C and P symmetries are violated; but if we interchange + with - , and also right with left, then it might seem that symmetry could be regained (called CP symmetry). However, CP symmetry is also violated in the K_L^o-decay, because there is an absolute rate difference between the final states $e^+ \pi^- \nu$ and $e^- \pi^+ \bar{\nu}$. As we shall see later in Chapter 15, from CP asymmetry and observed amplitudes of various K-decays we can deduce the asymmetry with respect to time-reversal T. At present, it appears that physical laws are not

symmetrical with respect to C, P, T, CP, PT and TC. Nevertheless, all indications are that the joint action of CPT (i.e., particle $\leftrightarrows$ antiparticle, right $\leftrightarrows$ left and past $\leftrightarrows$ future) remains a good symmetry.

Chapter 10

U_1 SYMMETRY AND P, C INVARIANCE

10.1 QED as an Example

The use of some of these symmetry operations in a quantum theory can best be illustrated by considering quantum electrodynamics (QED). In this section we discuss the conservation of lepton number and the invariance under parity P and particle-antiparticle conjugation C. As in Chapter 6, we adopt the Coulomb gauge. Hence the independent generalized coordinates are the transverse field $\vec{A} = \vec{A}^{tr}$ and the electron field ψ. Their conjugate momenta $\vec{\Pi}$ and $\mathcal{P}$ are given by (6.21)-(6.22). In accordance with (6.30), we have

$$\vec{\nabla} \cdot \vec{A} = 0 \qquad \text{and} \qquad \vec{\nabla} \cdot \vec{\Pi} = 0 \ .$$

Theorem. The QED Hamiltonian H, given by (6.55), is invariant under the following unitary transformations:

1. $$e^{iL\theta} H e^{-iL\theta} = H \tag{10.1}$$

in which θ is any real number and L is the lepton number operator defined by the normal product

$$L \equiv \int : \psi^{\dagger}(x) \psi(x) : d^3r \tag{10.2}$$

where, as before, $x_\mu = (\vec{r}, it)$.

2. $$C H C^{\dagger} = H \tag{10.3}$$

where C is the particle-antiparticle conjugation operator which satisfies

$$C\vec{A}(x)\,C^{\dagger} = -\vec{A}(x) \quad , \quad C\vec{\Pi}(x)\,C^{\dagger} = -\vec{\Pi}(x) \tag{10.4}$$

and

$$C\psi(x)\,C^{\dagger} = \eta_c\,\psi^c(x) \tag{10.5}$$

with η_c as a constant phase factor, $|\eta_c| = 1$, and the components of ψ^c given by, in the notation of Chapter 3,

$$\psi_a^c(x) \equiv (\gamma_2)_{a\lambda}\,\psi_\lambda^{\dagger}(x) \quad . \tag{10.6}$$

3. $$P\,H\,P^{\dagger} = H \tag{10.7}$$

where P is the space inversion (i.e., parity) operator which satisfies

$$P\vec{A}(\vec{r},t)\,P^{\dagger} = -\vec{A}(-\vec{r},t) \quad , \quad P\vec{\Pi}(\vec{r},t)\,P^{\dagger} = -\vec{\Pi}(-\vec{r},t) \quad , \tag{10.8}$$

$$P\psi(\vec{r},t)\,P^{\dagger} = \eta_p\,\gamma_4\,\psi(-\vec{r},t) \quad , \tag{10.9}$$

with η_p as a constant phase factor $|\eta_p| = 1$.

As we shall see, the operators $e^{iL\theta}$, C and P are all unitary. Furthermore, because of (10.1), (10.3) and (10.7), they are all t-independent.

Proof. 1. We first consider the commutator between L and $\psi(x)$, which on account of (3.24) and (10.2), is given by

$$[L, \psi(x)] = -\psi(x) \quad . \tag{10.10}$$

Next, we introduce

$$\psi_\theta(x) \equiv e^{iL\theta}\,\psi(x)\,e^{-iL\theta} \tag{10.11}$$

whose derivative is

$$\frac{\partial\psi_\theta(x)}{\partial\theta} = e^{iL\theta}\,i[L, \psi(x)]\,e^{-iL\theta} = -i\psi_\theta(x) \quad . \tag{10.12}$$

Equation (10.12) can be readily integrated. We find

$$\psi_\theta(x) = e^{-i\theta}\,\psi_{\theta=0}(x) = e^{-i\theta}\,\psi(x) \quad ,$$

which leads to

$$e^{iL\theta}\,\psi(x)\,e^{-iL\theta} = e^{-i\theta}\,\psi(x) \quad . \tag{10.13}$$

From the definition (10.2) we see that L is a Hermitian and therefore $e^{iL\theta}$ is unitary. Consequently, the Hermitian conjugate of (10.13) is

$$e^{iL\theta}\,\psi^\dagger(x)\,e^{-iL\theta} = e^{i\theta}\,\psi^\dagger(x) \quad . \tag{10.14}$$

Let Γ be any 4×4 matrices, such as γ_μ, $i\gamma_4\gamma_\mu$, $\cdots$, introduced in Chapter 3. It then follows that

$$\begin{aligned} e^{iL\theta}\,\psi^\dagger\Gamma\psi\,e^{-iL\theta} &= e^{iL\theta}\,\psi^\dagger\,e^{-iL\theta}\,\Gamma\,e^{iL\theta}\,\psi\,e^{-iL\theta} \\ &= \psi^\dagger e^{i\theta}\,\Gamma\,e^{-i\theta}\,\psi = \psi^\dagger\Gamma\psi \end{aligned} \tag{10.15}$$

in which we have used the fact that the matrix element of Γ is a c. number and therefore commutes with the Hilbert space operator L. In QED the electromagnetic current is $j_\mu = i : \psi^\dagger\gamma_4\gamma_\mu\psi :$. Hence,

$$e^{iL\theta}\,j_\mu(x)\,e^{-iL\theta} = j_\mu(x) \quad . \tag{10.16}$$

Furthermore, since L depends only on the fermion field, we have

$$e^{iL\theta}\,\vec{A}(x)\,e^{-iL\theta} = \vec{A}(x) \quad \text{and} \quad e^{iL\theta}\,\vec{\Pi}(x)\,e^{-iL\theta} = \vec{\Pi}(x) \quad . \tag{10.17}$$

Because $j_4 = i\rho$, and $A_0(\vec{r}, t)$ is given by (6.16), we also have

$$e^{iL\theta}\,A_0(x)\,e^{-iL\theta} = A_0(x) \quad , \tag{10.18}$$

and therefore

$$e^{iL\theta}\,j_\mu A_\mu\,e^{-iL\theta} = j_\mu A_\mu \quad .$$

Equation (10.1) now follows. By differentiating (10.1) and setting $\theta = 0$, we obtain

$$[L, H] = 0 \quad , \tag{10.19}$$

which implies L is independent of time, i.e., L is conserved.

From the Fourier expansion (3.32),

$$\psi(x) = \sum_{\vec{p},s} \frac{1}{\sqrt{\Omega}} \left(a_{\vec{p},s}\, u_{\vec{p},s}\, e^{i\vec{p}\cdot\vec{r}} + b^{\dagger}_{\vec{p},s}\, v_{\vec{p},s}\, e^{-i\vec{p}\cdot\vec{r}}\right), \quad (10.20)$$

we have

$$e^{iL\theta}\,\psi(x)\,e^{-iL\theta} = \sum_{\vec{p},s} \frac{1}{\sqrt{\Omega}} \left(e^{iL\theta} a_{\vec{p},s} e^{-iL\theta} u_{\vec{p},s} e^{i\vec{p}\cdot\vec{r}} + e^{iL\theta} b^{\dagger}_{\vec{p},s} e^{-iL\theta} v_{\vec{p},s} e^{-i\vec{p}\cdot\vec{r}}\right) \quad (10.21)$$

which, together with (10.13), leads to

$$e^{iL\theta} b^{\dagger}_{\vec{p},s} e^{-iL\theta} = e^{-i\theta} b^{\dagger}_{\vec{p},s} \quad (10.22)$$

and

$$e^{iL\theta} a_{\vec{p},s} e^{-iL\theta} = e^{-i\theta} a_{\vec{p},s} \ ,$$

whose Hermitian conjugate is

$$e^{iL\theta} a^{\dagger}_{\vec{p},s} e^{-iL\theta} = e^{i\theta} a^{\dagger}_{\vec{p},s} \ . \quad (10.23)$$

By using (3.32)-(3.33) we can readily verify that

$$L = \sum_{\vec{p},s} \left(a^{\dagger}_{\vec{p},s}\, a_{\vec{p},s} - b^{\dagger}_{\vec{p},s}\, b_{\vec{p},s}\right) \ . \quad (10.24)$$

As in (3.40), the Hilbert space of QED is spanned by the set of basis vectors

$$|\,0>, \quad a^{\dagger}_{\vec{p},s}\,|\,0>, \quad b^{\dagger}_{\vec{p},s}\,|\,0>, \quad a^{\dagger}_{\vec{p},s}\,|\,0>, \cdots$$

$$|\,n_{-}e^{-}, n_{+}e^{+}, n_{\gamma}\gamma> = \prod_{i=1}^{n_-}\prod_{j=1}^{n_+}\prod_{k=1}^{n_\gamma} a^{\dagger}_{\vec{p}_i,s_i}\, b^{\dagger}_{\vec{p}_j,s_j}\, a^{\dagger}_{\vec{p}_k,s_k}\,|\,0>, \quad (10.25)$$

$\cdots$, where $a^{\dagger}_{\vec{p},s}$ is the photon creation operator given by (6.36). Because of (3.41) and (10.24),

$$L\,|\,0> = 0 \ . \quad (10.26)$$

Furthermore, by following the discussions given in Section 3.5, we see that

$$L \,|\, n_- e^-, n_+ e^+, n_\gamma \gamma > \;=\; (n_- - n_+) \,|\, n_- e^-, n_+ e^+, n_\gamma \gamma > \;, \tag{10.27}$$

and therefore

$$e^{iL\theta} \,|\, n_- e^-, n_+ e^+, n_\gamma \gamma > \;=\; e^{i(n_- - n_+)\theta} \,|\, n_- e^-, n_+ e^+, n_\gamma \gamma > \;, \tag{10.28}$$

which can also be established directly by using (10.22)-(10.23); then by differentiating (10.28) with respect to θ and setting $\theta = 0$, we arrive back at (10.27).

2. Next, we consider the operator C defined by (10.4)-(10.5). From the Fourier expansion it follows that

$$C\psi(x)C^\dagger = \sum_{\vec{p},s} \frac{1}{\sqrt{\Omega}} \left(C a_{\vec{p},s} C^\dagger u_{\vec{p},s} e^{i\vec{p}\cdot\vec{r}} + C b^\dagger_{\vec{p},s} C^\dagger v_{\vec{p},s} e^{-i\vec{p}\cdot\vec{r}} \right), \tag{10.29}$$

in which we have used the property that $C^\dagger$ is a Hilbert space operator and therefore commutes with $e^{i\vec{p}\cdot\vec{r}}$ and the c. number spinors $u_{\vec{p},s}$ and $v_{\vec{p},s}$. In terms of the same expansion, (10.6) can be written as

$$\psi^c_a(x) = (\gamma_2)_{a\lambda} \psi^\dagger_\lambda(x) = \sum_{\vec{p},s} \frac{1}{\sqrt{\Omega}}$$

$$\cdot \left[a_{\vec{p},s} (\gamma_2)_{a\lambda} (u^*_{\vec{p},s})_\lambda e^{-i\vec{p}\cdot\vec{r}} + b_{\vec{p},s} (\gamma_2)_{a\lambda} (v^*_{\vec{p},s})_\lambda e^{i\vec{p}\cdot\vec{r}} \right]$$

which, because of (3.81), leads to

$$\psi^c(x) = \sum_{\vec{p},s} \frac{1}{\sqrt{\Omega}} \left(b_{\vec{p},s} u_{\vec{p},s} e^{i\vec{p}\cdot\vec{r}} + a^\dagger_{\vec{p},s} v_{\vec{p},s} e^{-i\vec{p}\cdot\vec{r}} \right) . \tag{10.30}$$

From (10.5) and by equating (10.29) with η_c times (10.30), we obtain

$$C a_{\vec{p},s} C^\dagger = \eta_c b_{\vec{p},s} \;, \tag{10.31}$$

$$C b^\dagger_{\vec{p},s} C^\dagger = \eta_c a^\dagger_{\vec{p},s} \;. \tag{10.32}$$

The Hermitan conjugate of (10.31) is

$$C\, a^{\dagger}_{\vec{p},s}\, C^{\dagger} = \eta_c^{*}\, b^{\dagger}_{\vec{p},s} \quad . \tag{10.33}$$

Likewise, by using (10.4) and the Fourier expansion (6.32)-(6.33) and (6.36), we can derive for the photon creation operator

$$C\, a^{\dagger}_{\vec{p},s}\, C^{\dagger} = -\, a^{\dagger}_{\vec{p},s} \quad . \tag{10.34}$$

That there indeed exists a unitary operator C which satisfies (10.32)-(10.34) is indicated by the formula given in Problem 10.2. Here we shall give a simpler and more direct proof. Let us consider a linear transformation in the Hilbert space spanned by (10.25):

$$\begin{aligned} &|0> \to |0>\,, \quad a^{\dagger}_{\vec{p},s}\,|0> \to \eta_c^{*}\, b^{\dagger}_{\vec{p},s}\,|0>\,, \\ &b^{\dagger}_{\vec{p},s}\,|0> \to \eta_c\, a^{\dagger}_{\vec{p},s}\,|0>\,, \quad a^{\dagger}_{\vec{p},s}\,|0> \to -\,a^{\dagger}_{\vec{p},s}\,|0>, \cdots \\ &|n_- e^-, n_+ e^+, n_\gamma \gamma> \to \eta_c^{*n_-}\, \eta_c^{n_+} (-)^{n_\gamma}\, |n_- e^+, n_+ e^-, n_\gamma \gamma> \,. \end{aligned} \tag{10.35}$$

The orthonormality of the complete set of basis vectors is clearly unchanged. Hence the transformation operator is unitary. Furthermore, one can easily see that it satisfies (10.32)-(10.34). Therefore, C exists and

$$C^{\dagger} C = 1 \quad . \tag{10.36}$$

In addition, in (10.35) we have fixed the overall phase factor of C by requiring $|0> \to |0>$; i.e.,

$$C\,|0> = |0> \quad . \tag{10.37}$$

Next, we shall establish the invariance of H under C. Since the free Hamiltonian H_0 of (6.55) can be written as

$$H_0 = \sum_{\vec{p},s=\pm\frac{1}{2}} E_p (a^{\dagger}_{\vec{p},s}\, a_{\vec{p},s} + b^{\dagger}_{\vec{p},s}\, b_{\vec{p},s}) + \sum_{\vec{k},s=\pm 1} \omega\, a^{\dagger}_{\vec{k},s}\, a_{\vec{k},s} \quad , \tag{10.38}$$

where $\omega = |\vec{k}|$, and $E_p = \sqrt{\vec{p}^{\,2} + m^2}$ with m the physical electron mass, from (10.32)-(10.34) one sees readily

$$C H_0 C^\dagger = H_0 \quad . \tag{10.39}$$

In accordance with our convention (3.7)-(3.11), $\gamma_2 = \rho_2 \sigma_2$ is real and symmetric. Hence the Hermitian conjugate of

$$C\psi_a C^\dagger = \eta_c (\gamma_2)_{a\lambda} \psi_\lambda^\dagger \tag{10.40}$$

is

$$C\psi_a^\dagger C^\dagger = \eta_c^* (\gamma_2)_{a\lambda} \psi_\lambda = \eta_c^* \psi_\lambda (\gamma_2)_{\lambda a} \quad . \tag{10.41}$$

The electromagnetic current operator is $j_\mu = i : \psi^\dagger \gamma_4 \gamma_\lambda \psi :$. By using (10.40)-(10.41), we find

$$\begin{aligned} C j_\mu C^\dagger &= i : C\psi_a^\dagger C^\dagger (\gamma_4 \gamma_\mu)_{a\beta} C\psi_\beta C^\dagger : \\ &= i : \psi_\lambda (\gamma_2)_{\lambda a} (\gamma_4 \gamma_\mu)_{a\beta} (\gamma_2)_{\beta\lambda'} \psi_{\lambda'}^\dagger : \\ &= -i : \psi_{\lambda'}^\dagger (\gamma_2 \gamma_4 \gamma_\mu \gamma_2)_{\lambda\lambda'} \psi_\lambda : \end{aligned} \tag{10.42}$$

where the minus sign is due to the exchange of the order of the fermion operators ψ_λ and $\psi_{\lambda'}^\dagger$, but with their anticommutator absent because of the normal product. Since γ_2, γ_4 are symmetric and γ_1, γ_3 antisymmetric, the transpose of $\gamma_2 \gamma_4 \gamma_\mu \gamma_2$ is given by

$$\widetilde{\gamma_2 \gamma_4 \gamma_\mu \gamma_2} = \gamma_2 \tilde{\gamma}_\mu \gamma_4 \gamma_2 = \gamma_4 \gamma_\mu \quad . \tag{10.43}$$

Substituting (10.43) into (10.42), we derive

$$C j_\mu C^\dagger = - j_\mu \quad . \tag{10.44}$$

Since $j_4 = i\rho$, it follows that

$$C\rho C^\dagger = - \rho \quad , \tag{10.45}$$

which, together with (6.16), yields

$$C A_0 C^\dagger = - A_0 \ . \tag{10.46}$$

Combining (10.4) with (10.46), we can write

$$C A_\mu C^\dagger = - A_\mu \ , \tag{10.47}$$

and that gives

$$C j_\mu A_\mu C^\dagger = j_\mu A_\mu$$

which establishes the particle-antiparticle conjugation symmetry of the electromagnetic interaction.

3. Lastly, we examine the parity operator P, defined by (10.8)-(10.9). Upon substituting the Fourier expansion (10.20) into (10.9), we find that its lefthand side becomes

$$P\psi(\vec{r}, t) P^\dagger = \sum_{\vec{p},s} (P a_{\vec{p},s} P^\dagger u_{\vec{p},s} e^{i\vec{p}\cdot\vec{r}} + P b^\dagger_{\vec{p},s} P^\dagger v_{\vec{p},s} e^{-i\vec{p}\cdot\vec{r}}) , \tag{10.48}$$

and its righthand side, on account of (3.82), can be written as

$$\begin{aligned} \eta_p \gamma_4 \psi(-\vec{r}, t) &= \sum_{\vec{p},s} \eta_p (a_{\vec{p},s} \gamma_4 u_{\vec{p},s} e^{-i\vec{p}\cdot\vec{r}} + b^\dagger_{\vec{p},s} \gamma_4 v_{\vec{p},s} e^{i\vec{p}\cdot\vec{r}}) \\ &= \sum_{\vec{p},s} \eta_p (a_{\vec{p},s} u_{-\vec{p},-s} e^{-i\vec{p}\cdot\vec{r}} - b^\dagger_{\vec{p},s} v_{-\vec{p},-s} e^{i\vec{p}\cdot\vec{r}}) \\ &= \sum_{\vec{p},s} \eta_p (a_{-\vec{p},-s} u_{\vec{p},s} e^{i\vec{p}\cdot\vec{r}} - b^\dagger_{-\vec{p},-s} v_{\vec{p},s} e^{-i\vec{p}\cdot\vec{r}}) . \end{aligned} \tag{10.49}$$

Hence, by equating (10.48) with (10.49), we obtain

$$P b^\dagger_{\vec{p},s} P^\dagger = - \eta_p b^\dagger_{-\vec{p},-s} \tag{10.50}$$

and

$$P a_{\vec{p},s} P^\dagger = \eta_p a_{-\vec{p},-s} \ , \tag{10.51}$$

whose Hermitian conjugate is

$$P a^\dagger_{\vec{p},s} P^\dagger = \eta_p^* a^\dagger_{-\vec{p},-s} \ . \tag{10.52}$$

Likewise, by using (10.8) and the Fourier expansion (6.32)-(6.33), we

can derive for the photon creation operator

$$P \vec{a}^{\dagger}_{\vec{p}} P^{\dagger} = - \vec{a}^{\dagger}_{-\vec{p}} \tag{10.53}$$

where, as in (6.34),

$$\vec{p} \cdot \vec{a}^{\dagger}_{\vec{p}} = 0 \quad . \tag{10.54}$$

For photons of a definite helicity, we may use the creation operator defined by (6.36). Equation (10.53) takes on the form

$$P\, a^{\dagger}_{\vec{p},s}\, P^{\dagger} = - a^{\dagger}_{-\vec{p},-s} \quad . \tag{10.55}$$

Next, we consider a linear transformation in the Hilbert space spanned by (10.25):

$$|0> \rightarrow |0>, \quad a^{\dagger}_{\vec{p},s}|0> \rightarrow \eta_p^{*}\, a^{\dagger}_{-\vec{p},-s}|0>,$$

$$b^{\dagger}_{\vec{p},s}|0> \rightarrow -\eta_p\, b^{\dagger}_{-\vec{p},-s}|0>, \quad a^{\dagger}_{\vec{p},s}|0> \rightarrow -a^{\dagger}_{-\vec{p},-s}|0>, \cdots \tag{10.56}$$

which, as in (10.35), preserves the orthonormality of the complete set of basis vectors, and is therefore unitary. One can readily see that the transformation operator satisfies (10.51)-(10.53). Therefore P exists and

$$P^{\dagger} P = 1 \quad ; \tag{10.57}$$

furthermore, because of the first expression in (10.56), we have fixed the overall phase factor of P by setting

$$P|0> = |0> \quad . \tag{10.58}$$

From (10.38), one observes that the free Hamiltonian is invariant under P ; i.e.

$$P H_0 P^{\dagger} = H_0 \quad . \tag{10.59}$$

As in (10.42), the transformation of the electromagnetic current

operator j_μ under P is given by

$$\begin{aligned} P\, j_\mu(\vec{r}, t)\, P^\dagger &= i : P\psi^\dagger(\vec{r}, t)\, P^\dagger \gamma_4 \gamma_\mu P\psi(\vec{r}, t)\, P^\dagger : \\ &= i : \psi^\dagger(-\vec{r}, t)\, \gamma_4 \gamma_4 \gamma_\mu \gamma_4 \psi(-\vec{r}, t) : \\ &= \begin{cases} -\, j_i(-\vec{r}, t) & \mu = i \neq 4 \\ +\, j_4(-\vec{r}, t) & \mu = 4 \end{cases} , \end{aligned} \tag{10.60}$$

which, together with (10.8), leads to

$$P\, \vec{j}(\vec{r}, t) \cdot \vec{A}(\vec{r}, t)\, P^\dagger = \vec{j}(-\vec{r}, t) \cdot \vec{A}(-\vec{r}, t) \quad . \tag{10.61}$$

Since $j_4 = i\rho$ and $\nabla^2 A_0 = -e\rho$, we have

$$P\, \rho(\vec{r}, t)\, A_0(\vec{r}, t)\, P^\dagger = \rho(-\vec{r}, t)\, A_0(-\vec{r}, t) \quad . \tag{10.62}$$

Furthermore, because of (10.9) and its Hermitian conjugate, we obtain

$$P\, \psi^\dagger(\vec{r}, t)\, \beta\, \psi(\vec{r}, t)\, P^\dagger = \psi^\dagger(-\vec{r}, t)\, \beta\, \psi(-\vec{r}, t) \quad . \tag{10.63}$$

Combining (10.61)-(10.63) with (6.28), we see that

$$P\, H_{int}\, P^\dagger = H_{int} \quad . \tag{10.64}$$

Equations (10.59) and (10.64) establish the invariance of H under P.

Remarks. Under the space inversion, the momentum $\vec{p}$ of a particle changes sign, but its spin $\vec{\sigma}$ does not, and therefore its helicity $s = \vec{\sigma} \cdot \hat{p} \to -s$, in accordance with (10.50)-(10.55). For a nonrelativistic particle, its spin is decoupled from its momentum; we may replace (3.27) by

$$\vec{\sigma} \cdot \hat{z} \begin{cases} u_{\vec{p}, \sigma_z} \\ v_{-\vec{p}, \sigma_z} \end{cases} = \sigma_z \begin{cases} u_{\vec{p}, \sigma_z} \\ -\, v_{-\vec{p}, \sigma_z} \end{cases} \tag{10.65}$$

where $\hat{z}$ is a unit vector along the z-axis, which can be arbitrarily chosen, and $\frac{1}{2}\sigma_z = \pm\frac{1}{2}$ is the spin of the particle along $\hat{z}$.

Correspondingly, (3.31) becomes

$$S_{\vec{p}}(t) = \sum_{\sigma_z=\pm 1} (a_{\vec{p},\sigma_z}(t)\, u_{\vec{p},\sigma_z} + b^{\dagger}_{-\vec{p},\sigma_z}(t)\, v_{-\vec{p},\sigma_z}) \,. \quad (10.66)$$

From (10.52)-(10.53), one can readily derive

$$P\, a^{\dagger}_{\vec{p},\sigma_z}\, P^{\dagger} = \eta_p^{*}\, a^{\dagger}_{-\vec{p},\sigma_z} \,,$$

$$P\, b^{\dagger}_{\vec{p},\sigma_z}\, P^{\dagger} = -\eta_p\, b^{\dagger}_{-\vec{p},\sigma_z} \,, \quad (10.67)$$

which shows explicitly that spin, being an angular momentum, is a pseudovector and therefore does not change sign under P.

10.2 Applications

1. Furry theorem

Let us consider a state of n photons

$$|\, n\gamma > \equiv \prod_{i=1}^{n} a^{\dagger}_{\vec{p}_i,s_i} \,|\, 0 > \,. \quad (10.68)$$

From (10.34) and (10.37) we have

$$C \,|\, n\gamma > = (-1)^n \,|\, n\gamma > \,. \quad (10.69)$$

Since the QED Hamiltonian is C-invariant, the $U(t, t_0)$ matrix defined by (5.20)-(5.21) satisfies

$$C\, U(t, t_0)\, C^{\dagger} = U(t, t_0) \,. \quad (10.70)$$

In the limits $t \to \infty$ and $t_0 \to -\infty$, $U(t, t_0)$ becomes the S-matrix. Hence,

$$C\, S\, C^{\dagger} = S \,. \quad (10.71)$$

We may consider the matrix element

$$< n'\gamma \,|\, S \,|\, n\gamma > \,, \quad (10.72)$$

which, because of (10.69) and (10.71), is equal to

$$\langle n'\gamma \mid C^{\dagger}SC \mid n\gamma \rangle = (-1)^{n+n'} \langle n'\gamma \mid S \mid n\gamma \rangle \quad , \qquad (10.73)$$

which must be zero if $n + n'$ is an odd number. Therefore, we establish the theorem

$$\text{even numbers of } \gamma \not\rightarrow \text{odd numbers of } \gamma \quad , \qquad (10.74)$$

valid to all orders of the electromagnetic interaction.

The above selection rule can be easily extended to include the strong interaction, provided that the strong-interaction Hamiltonian H_{st} is also assumed to be C-invariant; i.e.

$$C H_{st} C^{\dagger} = H_{st} \quad . \qquad (10.75)$$

Selection rule (10.74) is then valid to all orders of the strong, as well as the electromagnetic, interaction.

A glance at the Table of Particle Properties at the end of this book tells one that the dominant decay mode of π^{o} is

$$\pi^{o} \rightarrow 2\gamma \quad . \qquad (10.76)$$

From assumption (10.75) and Eq. (10.69), we conclude π^{o} is of $C = +1$, and therefore

$$\pi^{o} \not\rightarrow \text{odd number of } \gamma \quad . \qquad (10.77)$$

The experimental upper bound of the branching ratio of $\pi^{o} \rightarrow 3\gamma$ is $< 5 \times 10^{-6}$, which confirms the above selection rule and in turn gives support to the correctness of our assumption that H_{st} is C-invariant.

Remarks. In the case of QED, C-invariance is established explicitly; in the case of the strong interaction, it is assumed. In either event, the deduction of selection rules, such as (10.74) and (10.77), from C-invariance can be made independently of the details of the Hamiltonian.

The close correlation between the general theoretical assumptions and their rigorous experimental consequences makes the study of symmetry particularly rewarding.

2. Positronium states

It is customary to label the positronium e^+e^- states as $^{2S+1}L_J$ where the quantum number L is the orbital angular momentum, J the total angular momentum and S the total spin, with $S = 0$ being the singlet and $S = 1$ the triplet. We first establish that such a state is of

$$P = -(-1)^L \quad \text{and} \quad C = (-1)^{L+S} \; . \tag{10.78}$$

Proof. Because the electromagnetic interaction is both C- and P-invariant, we can adiabatically reduce the electromagnetic coupling constant e^2 without altering the C and P quantum numbers of the positronium state $^{2S+1}L_J$. In the limit of very small e^2, the non-relativistic description becomes sufficient, and therefore we can neglect the coupling between spin and orbital angular momentum as well as the presence of virtual photons. The state vector is then given by

$$|\, ^{2S+1}L_J > = \sum_{m, \sigma_z, \sigma_z'} \int d^3p \, Y_{LM}(\hat{p}) \, \chi_m(\sigma_z, \sigma_z') \, C_{Mm}(p) \cdot a^\dagger_{\vec{p}, \sigma_z} \, b^\dagger_{-\vec{p}, \sigma_z'} \, |\, 0 > \tag{10.79}$$

where, as in (10.65)-(10.66), subscripts σ_z and σ_z' can be ± 1; they denote the spin components of e^- and e^+ along the z-axis, with the spin function $\chi_m(\sigma_z, \sigma_z')$ antisymmetric for the singlet, but symmetric for the triplet,

$$\chi_m(\sigma_z, \sigma_z') = (-1)^{S+1} \chi_m(\sigma_z', \sigma_z) \; . \tag{10.80}$$

Quite often in the literature one writes for the triplet state, $\chi_m = \uparrow_-\uparrow_+$

for $m = 1$, $(\uparrow_-\uparrow_+ + \downarrow_-\downarrow_+)/\sqrt{2}$ for $m = 0$, and $\downarrow_-\downarrow_+$ for $m = -1$; for the singlet state one has only $m = 0$ and the corresponding $X_m = (\uparrow_-\downarrow_+ - \downarrow_-\uparrow_+)/\sqrt{2}$, where the subscript $-$ is for e^- and $+$ for e^+. In (10.79) the momenta of e^- and e^+ are $\vec{p}$ and $-\vec{p}$ respectively, $\hat{p} = \vec{p}/p$ with $p = |\vec{p}|$, Y_{LM} denotes the spherical harmonics given by (1.38)–(1.40) which satisfies

$$Y_{LM}(-\hat{p}) = (-1)^L Y_{LM}(\hat{p}) \quad . \tag{10.81}$$

The z-component of the total angular momentum of the state is $J_z = M + m$; $C_{Mm}(p)$ denotes the appropriate Clebsch–Gordon coefficient multiplied by the radial function, which is independent of the direction of $\vec{p}$.

We now apply the Hilbert space operator P onto the state vector (10.79):

$$P\,|\,^{2S+1}L_J> = \sum_{m,\sigma_z,\sigma_z'} \int d^3p\, Y_{LM}(\hat{p})\, X_m(\sigma_z, \sigma_z')\, C_{Mm}(p) \cdot P\, a^\dagger_{\vec{p},\sigma_z}\, b^\dagger_{-\vec{p},\sigma_z'} \,|\, 0> \quad . \tag{10.82}$$

Because of (10.58) and (10.67),

$$P\, a^\dagger_{\vec{p},\sigma_z}\, b^\dagger_{-\vec{p},\sigma_z'} \,|\, 0> = P\, a^\dagger_{\vec{p},\sigma_z}\, P^\dagger P\, b^\dagger_{-\vec{p},\sigma_z'}\, P^\dagger P \,|\, 0> = -\, a^\dagger_{-\vec{p},\sigma_z}\, b^\dagger_{\vec{p},\sigma_z'} \,|\, 0> \tag{10.83}$$

which, together with (10.82), gives

$$P\,|\,^{2S+1}L_J> = -\sum_{m,\sigma_z,\sigma_z'} \int d^3p\, Y_{LM}(\hat{p})\, X_m(\sigma_z, \sigma_z')\, C_{Mm}(p) \cdot a^\dagger_{-\vec{p},\sigma_z}\, b^\dagger_{\vec{p},\sigma_z'} \,|\, 0> \quad . \tag{10.84}$$

By re-labeling $-\vec{p}$ as $\vec{p}$ and by using (10.81), we can rewrite (10.84):

$$P\,|^{2S+1}L_J> = -\sum_{m,\sigma_z,\sigma_z'} \int d^3p\, Y_{LM}(-\hat{p})\, \chi_m(\sigma_z, \sigma_z')\, C_{Mm}(p) \cdot a^\dagger_{\vec{p},\sigma_z}\, b^\dagger_{-\vec{p},\sigma_z'}\,|0>$$

$$= -(-1)^L\,|^{2S+1}L_J> \,, \qquad (10.85)$$

which gives the first equation in (10.78).

By using (10.32)-(10.33) and (10.37), we have

$$C\,a^\dagger_{\vec{p},\sigma_z}\,b^\dagger_{-\vec{p},\sigma_z'}\,|0> = C\,a^\dagger_{\vec{p},\sigma_z}\,C^\dagger C\,b_{-\vec{p},\sigma_z'}\,C^\dagger C\,|0>$$

$$= b^\dagger_{\vec{p},\sigma_z}\,a^\dagger_{-\vec{p},\sigma_z'}\,|0>$$

$$= -\,a^\dagger_{-\vec{p},\sigma_z'}\,b^\dagger_{\vec{p},\sigma_z}\,|0> \,. \qquad (10.86)$$

Hence, the application of C to the state vector (10.79) gives

$$C\,|^{2S+1}L_J> = \sum_{m,\sigma_z,\sigma_z'} \int d^3p\, Y_{LM}(\hat{p})\, \chi_m(\sigma_z, \sigma_z')\, C_{Mm}(p) \cdot C\,a^\dagger_{\vec{p},\sigma_z}\,b^\dagger_{-\vec{p},\sigma_z'}\,|0>$$

$$= -\sum_{m,\sigma_z,\sigma_z'} \int d^3p\, Y_{LM}(\hat{p})\, \chi_m(\sigma_z, \sigma_z')\, C_{Mm}(p) \cdot a^\dagger_{-\vec{p},\sigma_z'}\,b^\dagger_{\vec{p},\sigma_z}\,|0>$$

$$= -\sum_{m,\sigma_z,\sigma_z'} \int d^3p\, Y_{LM}(-\hat{p})\, \chi_m(\sigma_z', \sigma_z)\, C_{Mm}(p) \cdot a^\dagger_{\vec{p},\sigma_z}\,b^\dagger_{-\vec{p},\sigma_z'}\,|0> \qquad (10.87)$$

which, because of (10.80) and (10.81), is

$$(-1)^{L+S}\,|^{2S+1}L_J>$$

and that establishes the second equation in (10.78).

Combining (10.78) with (10.69), we derive the selection rules:

states with even $L+S \not\to$ odd γ , (10.88)

which applies to, e.g., $^1S_0$, $^3P_0$, $^3P_1$, $^3P_2$, ⋯ states;

states with odd $L+S \not\to$ even γ , (10.89)

which applies to, e.g., $^3S_1$, $^1P_1$, ⋯ states.

Remarks. There is a simple method for remembering the two equations in (10.78). In accordance with (10.67), the relative parity between e^+ and e^- is -1 ; under space inversion, spin is unchanged, $\vec{p} \to -\vec{p}$ and therefore $Y_{LM} \to (-1)^L Y_{LM}$. Together they give $P = -(-1)^L$.

Under C, we exchange e^+ and e^-; this gives a factor $(-1)^L$ due to the orbital angular momentum, a factor $(-1)^{S+1}$ due to the spin part and another -1 due to Fermi statistics. Together they give $C = (-1)^L (-1)^{S+1} (-1) = (-1)^{L+S}$.

3. Decay of a spin-0 particle $\to 2\gamma$

Let us consider the rest frame of the spin-0 particle. In its 2γ final state, the momenta of these two photons are equal; each γ carries a polarization direction which can be defined to be that of its electric field. We shall now establish that:

for a scalar particle $(P=+1)$, the polarization directions of these two γ's are parallel, (10.90)

for a pseudoscalar particle $(P=-1)$, the two polarization directions are perpendicular. (10.91)

Proof. Any 2γ state in its center-of-mass system can be written as

$$| 2\gamma > = \int d^3p \, X_{ij}(\vec{p}) \, a_i^\dagger(\vec{p}) \, a_j^\dagger(-\vec{p}) \, | 0 > \ , \qquad (10.92)$$

where $a_i^\dagger(\vec{p})$ denotes the i^{th} component of the photon creation

operator $\vec{a}^{\dagger}_{\vec{p}}$ in (6.32)-(6.33). From (6.34) it follows that

$$p_i \, a_i^{\dagger}(\vec{p}) = 0 \tag{10.93}$$

where, as before, all repeated indices are to be summed over. For (10.92) to be of spin 0, the function $\chi_{ij}(\vec{p})$ must be a tensor of second rank under space rotation. Since it depends only on a single vector $\vec{p}$, the most general form of $\chi_{ij}(\vec{p})$ is

$$\chi_{ij}(\vec{p}) = A\delta_{ij} + B\,\epsilon_{ijk}\,p_k + C\,p_i\,p_j \ , \tag{10.94}$$

where δ_{ij} is the Kronecker symbol and

$$\epsilon_{ijk} = \begin{cases} +1 & \text{if } ijk \text{ is an even permutation of } 1, 2, 3, \\ -1 & \text{if } ijk \text{ is an odd permutation of } 1, 2, 3, \\ 0 & \text{otherwise.} \end{cases} \tag{10.95}$$

A, B, C are functions of $|\vec{p}|$. Because of (10.93), the term $Cp_i p_j$ in (10.94) makes no contribution to the $|2\gamma>$ state of (10.92); $\chi_{ij}(\vec{p})$ can be simplified to

$$\chi_{ij}(\vec{p}) = A\,\delta_{ij} + B\,\epsilon_{ijk}\,p_k \ . \tag{10.96}$$

On account of (10.53), we have in our present notation

$$P\,a_i^{\dagger}(\vec{p})\,P^{\dagger} = -\,a_i^{\dagger}(-\vec{p}) \ . \tag{10.97}$$

Hence,

$$\begin{aligned} P\,|\,2\gamma> &= \int d^3p\,\chi_{ij}(\vec{p})\,P\,a_i^{\dagger}(\vec{p})\,P^{\dagger}\,P\,a_j^{\dagger}(-\vec{p})\,P^{\dagger}P\,|\,0> \\ &= \int d^3p\,\chi_{ij}(\vec{p})\,a_i^{\dagger}(-\vec{p})\,a_j^{\dagger}(\vec{p})\,|\,0> \\ &= \int d^3p\,\chi_{ij}(-\vec{p})\,a_i^{\dagger}(\vec{p})\,a_j^{\dagger}(-\vec{p})\,|\,0> \ , \end{aligned} \tag{10.98}$$

which means that for states with $P = \pm 1$,

$$\chi_{ij}(-\vec{p}) = \pm\chi_{ij}(\vec{p}) \ .$$

Thus we obtain

$$\chi_{ij}(\vec{p}) = \begin{cases} A\,\delta_{ij} & P = 1 \ , \\ B\,\epsilon_{ijk}\,p_k & P = -1 \ . \end{cases} \tag{10.99}$$

According to the Fourier expansion (6.33), the vector direction of $\vec{a}^{\dagger}$ is the corresponding direction of $\vec{E}$. Equation (10.99) shows that the polarizations of two photons are parallel when $P = +1$, and perpendicular when $P = -1$, which establishes (10.90)-(10.91).

Assuming that π^o is of spin 0, we may use the final polarization directions of the two photons in

$$\pi^o \rightarrow 2\gamma$$

to determine the parity of π^o. In the decay

$$\pi^o \rightarrow \gamma_a + \gamma_b \rightarrow (e_a^+ + e_a^-) + (e_b^+ + e_b^-) \ ,$$

the plane determined by the pair $e_a^+ e_a^-$ contains the electric field vector of the parent photon γ_a (and an identical relation between $e_b^+ e_b^-$ and γ_b). Via the relative orientations of the two planes, determined by $e_a^+ e_a^-$ and $e_b^+ e_b^-$, Chinovsky and Steinberger* were able to establish that π^o is a pseudoscalar; i.e., $P_{\pi^o} = -1$. Again, this conclusion is valid so long as both the strong and electromagnetic interactions are P-invariant.

Remarks. The same results (10.90)-(10.91) can also be derived phenomenologically. We may represent the parent meson by a spin-0 field ϕ, and consider an effective Lagrangian density $\mathcal{L}_{eff}$ between ϕ and the electromagnetic fields of the 2γ, which will be denoted by a and b respectively. From Lorentz-invariance, we find that if ϕ is a scalar, then

---

* W. Chinovsky and J. Steinberger, Phys. Rev. 95, 1561 (1954). See also the references cited therein.

$$\mathcal{L}_{eff} \propto \phi F_{\mu\nu}(a) F_{\mu\nu}(b) \propto \phi [\vec{E}(a)\cdot\vec{E}(b) - \vec{B}(a)\cdot\vec{B}(b)] \ , \quad (10.100)$$

and if ϕ is a pseudoscalar,

$$\mathcal{L}_{eff} \propto \phi \epsilon_{\mu\nu\lambda\delta} F_{\mu\nu}(a) F_{\lambda\delta}(b) \propto \phi [\vec{E}(a)\cdot\vec{B}(b) + \vec{B}(a)\cdot\vec{E}(b)] \quad (10.101)$$

where

$$\epsilon_{\mu\nu\lambda\delta} = \begin{cases} +1 & \text{if } \mu\nu\lambda\delta \text{ is an even permutation of } 1,2,3,4, \\ -1 & \text{if } \mu\nu\lambda\delta \text{ is an odd permutation of } 1,2,3,4, \\ 0 & \text{otherwise.} \end{cases} \quad (10.102)$$

It is clear that (10.100), (10.101) give (10.90), (10.91) respectively.

4. Spin-1 particle $\not\to 2\gamma$

Let us try to construct a 2γ state with a total angular momentum $J = 1$ in the center-of-mass system. Such a state must transform like a vector under space rotation; like the component v_i of a vector $\vec{v}$, it must carry a similar index $i = 1,2,3$ making a total of three $J = 1$ states. We may therefore write

$$| 2\gamma >_i = \int d^3p \, \chi_{ijk}(\vec{p}) \, a_j^\dagger(\vec{p}) \, a_k^\dagger(-\vec{p}) \, | 0 > \quad (10.103)$$

where $\chi_{ijk}(\vec{p})$ is a tensor of third rank under space rotation. Since χ_{ijk} depends only on the single vector $\vec{p}$, its most general form is given by

$$\begin{aligned} \chi_{ijk}(\vec{p}) = {} & A \epsilon_{ijk} + B p_i \delta_{jk} + C p_j \delta_{ik} + D p_k \delta_{ij} \\ & + B' p_i \epsilon_{jk\ell} p_\ell + C' p_j \epsilon_{ki\ell} p_\ell + D' p_k \epsilon_{ij\ell} p_\ell \\ & + E p_i p_j p_k \end{aligned} \quad (10.104)$$

where $A, B, \cdots, D', E$ are functions only of $|\vec{p}|$. Because $p_i \, a_i^\dagger(\vec{p}) = 0$, in the above expression the C, D, C', D and E terms do not contribute to the 2γ states of (10.103). The above expression can be simplified to

$$\chi_{ijk}(\vec{p}) = A\,\epsilon_{ijk} + Bp_i\delta_{jk} + B'p_i\,\epsilon_{jk\ell}\,p_\ell \qquad (10.105)$$

which satisfies

$$\chi_{ijk}(\vec{p}) = -\chi_{ikj}(-\vec{p}) \quad . \qquad (10.106)$$

On the other hand, because the photons obey boson statistics, we may commute the two creation operators in (10.103):

$$|\,2\gamma\rangle_i = \int d^3p\,\chi_{ijk}(\vec{p})\,a_k^\dagger(-\vec{p})\,a_j^\dagger(\vec{p})\,|\,0\rangle$$

which, after a re-labeling of $\vec{p}$ as $-\vec{p}$ and j, k as k, j, becomes

$$|\,2\gamma\rangle_i = \int d^3p\,\chi_{ikj}(-\vec{p})\,a_j^\dagger(\vec{p})\,a_k^\dagger(-\vec{p})\,|\,0\rangle \quad . \qquad (10.107)$$

By adding (10.103) to (10.107) and by using (10.106), we obtain

$$\begin{aligned} |\,2\gamma\rangle_i &= \int d^3p\,\tfrac{1}{2}[\chi_{ijk}(\vec{p}) + \chi_{ikj}(-\vec{p})]\,a_j^\dagger(\vec{p})\,a_k^\dagger(-\vec{p})\,|\,0\rangle \\ &= 0 \quad . \end{aligned} \qquad (10.108)$$

That means it is not possible to put 2γ in a total angular momentum $J = 1$ state. Hence, a spin-1 particle $\not\to 2\gamma$. In the decay of positronium, all states with $J = 1$ must obey the selection rule

$$^{2S+1}L_{J=1} \not\to 2\gamma \quad . \qquad (10.109)$$

However, such states can decay into 3γ,

$$^{2S+1}L_{J=1} \to 3\gamma \quad . \qquad (10.110)$$

Likewise, from $\pi^o \to 2\gamma$ one knows that the spin of $\pi^o \neq 1$. Later on, in Chapter 13, we shall show that the pion spin is indeed 0.

10.3 General Discussion

Let us consider a general system whose Hamiltonian

$$H = H_0 + H_{int}$$

is invariant under a unitary transformation $\mathcal{S}$; i.e.,

$$\mathcal{S} H_0 \mathcal{S}^\dagger = H_0 \ , \quad \mathcal{S} H_{int} \mathcal{S}^\dagger = H_{int} \tag{10.111}$$

and

$$\mathcal{S}^\dagger \mathcal{S} = 1 \ . \tag{10.112}$$

For example, $\mathcal{S}$ can be $e^{iL\theta}$, or C, or P. From (10.111)–(10.112), it follows that

$$\mathcal{S} H_0 = H_0 \mathcal{S} \quad \text{and} \quad \mathcal{S} H_{int} = H_{int} \mathcal{S} \tag{10.113}$$

which, on account of (5.1) and (5.6), implies that $\mathcal{S}$ is independent of t in either the Heisenberg or the interaction representation. Furthermore, in the interaction representation the $U(t, t_0)$ matrix, defined by (5.20)-(5.21), commutes with $\mathcal{S}$,

$$[\mathcal{S}, U(t, t_0)] = 0 \ . \tag{10.114}$$

We may denote the eigenstate of the free Hamiltonian H_0 by $| n >_{free}$,

$$H_0 | n >_{free} = E_n | n >_{free} \ . \tag{10.115}$$

In accordance with (6.58) and (6.60), there exists a corresponding eigenstate $| n >_{phys}$ of the total Hamiltonian, related to $| n >_{free}$ by

$$| n >_{phys} \equiv U(t, -\infty) | n >_{free} \ . \tag{10.116}$$

From (10.113), the matrices H_0 and $\mathcal{S}$ can be diagonalized simultaneously. Thus, in (10.115) the state $| n >_{free}$ can be chosen to be

$$\mathcal{S} | n >_{free} = s | n >_{free}$$

where s is the eigenvalue. Because of (10.114) and (10.116), we have

$$\mathcal{S} | n >_{phys} = U(t, -\infty) \mathcal{S} | n >_{free} = s | n >_{phys} \ .$$

Hence, independently of the details of H_{int}, the invariance assumption

(10.111) implies that the transformation properties of the physical state $|\, n >_{phys}$ are completely determined by those of the corresponding free state $|\, n >_{free}$, and that greatly simplifies the mathematical analysis.

10.4 Baryon Number and Lepton Number

To each physical single-particle state we assign the following eigenvalue to the baryon-number operator N:

$N = 1$ for a single-baryon state, e.g., $p, n, \Lambda, \cdots,$

$N = -1$ for a single-antibaryon state, e.g., $\bar{p}, \bar{n}, \bar{\Lambda}, \cdots,$

$N = 0$ for all other single-particle states, e.g., $e^{\pm}, \pi^{\pm}, \pi^{o}, \gamma, \cdots.$ (10.117)

For a multiparticle state, the baryon number is given by the corresponding algebraic sum, $\sum N_i$. Hadrons with $N = 0$ are called mesons, otherwise baryons or antibaryons.

By definition, N is a Hermitian operator since its eigenvalues are all real numbers (in fact, integers). Let θ be a real number. The operator

$$U = e^{iN\theta} \tag{10.118}$$

is unitary. The assumption that the total Hamiltonian H is invariant under U:

$$U H U^{\dagger} = H \tag{10.119}$$

insures that

$$[N, H] = 0 \quad , \tag{10.120}$$

as can be readily verified by differentiating (10.119) with respect to θ. Thus, N is conserved. The converse is also true; i.e. (10.120) implies (10.119). From (10.118), we have

$$U\,|\,e> = |\,e>\ , \qquad U\,|\,p> = e^{i\theta}\,|\,p>\ ,$$

$$U\,|\,n> = e^{i\theta}\,|\,n>\ , \qquad U\,|\,2p> = e^{2i\theta}\,|\,2p>\ , \ \ldots\ .$$

If N is conserved, then since p is the lowest mass state with $N = 1$, it must be stable; e.g.

$$p \not\to e^+ + \gamma \qquad \text{and} \qquad p \not\to \mu^+ + \gamma\ .$$

Thus, the stability of the proton is tied to the conservation of baryon number.

Likewise, we can assign the e-lepton number L_e and the μ-lepton number L_μ:

	μ^-	μ^+	ν_μ	$\bar{\nu}_\mu$	e^-	e^+	ν_e	$\bar{\nu}_e$	other
L_μ	+1	-1	+1	-1	0	0	0	0	0
L_e	0	0	0	0	+1	-1	+1	-1	0

The lepton number of a multiparticle state is again the algebraic sum of the lepton numbers of its constituents. Hence, conservation of L_μ and L_e implies that, e.g. in the decay of π^+, depending on the charged lepton in the final state, the neutrino can be either ν_μ or ν_e:

$$\pi^+ \to \begin{cases} \mu^+ + \nu_\mu & (10.121) \\ e^+ + \nu_e & (10.122) \end{cases}\ .$$

Likewise, in the decay of its antiparticle π^- we have

$$\pi^- \to \begin{cases} \mu^- + \bar{\nu}_\mu & (10.123) \\ e^- + \bar{\nu}_e & (10.124) \end{cases}\ .$$

If one wishes, one may regard (10.121)-(10.124) as the definition of these neutrinos and antineutrinos. Because the strong interaction can cause virtual transitions $p \rightleftarrows n + \pi^+$ and $n \rightleftarrows p + \pi^-$, in μ^--capture and in β-decay, we must have the same kinds of neutrinos and antineutrinos,

$$\mu^- + p \rightarrow n + \nu_\mu \qquad \text{and} \qquad n \rightarrow p + e^- + \bar{\nu}_e \ . \tag{10.125}$$

Similarly, in μ-decay, we have

$$\mu^- \rightarrow e^- + \nu_\mu + \bar{\nu}_e \qquad \text{and} \qquad \mu^+ \rightarrow e^+ + \bar{\nu}_\mu + \nu_e \ .$$

The conservation of L_μ and L_e leads then to the following selection rules:

$$\mu^- \not\rightarrow e^- + \gamma \ , \tag{10.126}$$

$$\nu_\mu + n \not\rightarrow e^- + p \ , \tag{10.127}$$

$$Z \not\rightarrow (Z+2) + e^- + e^- \ . \tag{10.128}$$

Both (10.126) and (10.127) support only the conservation of the difference $L_\mu - L_e$; the absence of neutrino-less double β-decay (10.128) gives direct evidence for the conservation of L_e. Together they provide the experimental proof of these two lepton number conservations. [See Section 21.1 for a discussion of L_τ.]

Historically, the two-neutrino hypothesis, $\nu_\mu \neq \nu_e$, was introduced to "explain" the extremely small upper bound of the branching ratio for the $\mu \rightarrow e + \gamma$ decay,

$$\frac{\text{rate}\ (\mu^- \rightarrow e^- + \gamma)}{\text{rate}\ (\mu^- \rightarrow e^- + \nu_\mu + \bar{\nu}_e)} < 10^{-8} \ ,$$

and the high-energy neutrino experiment was suggested,* in part

* T.D. Lee and C.N. Yang, Phys. Rev. Lett. 4, 307 (1960); M. Schwartz, ibid. 306.

because of the search for ways to substantiate this hypothesis. By comparing the rate for reaction (10.127) with the allowed process $\nu_\mu + n \to \mu^- + p$, Lederman, Schwartz, Steinberger and their collaborators* were able to establish the validity of the two-neutrino hypothesis in 1962 by using high-energy neutrinos; this in turn helped to shape the present massive style of doing high-energy experimental physics.

In addition to the conservation of baryon number N and lepton numbers L_μ and L_e, we also have the familiar conservation of electric charge Q. Unlike the parity P and the particle-antiparticle conjugation C, which are multiplicatively conserved, N, L_μ, L_e and Q are all additively conserved. The unit of electric charge is a measurable quantity, as evidenced by the well-known fine-structure constant $\alpha \cong \frac{1}{137}$. In contrast, the units of N, L_μ and L_e are arbitrarily chosen. This is because while Q gives rise to the Coulomb field, so far as we know neither N nor L_μ, L_e are the sources of physical fields. This disparity is perhaps a very deep one.** It has led people to speculate that probably conservation of electric charge is truly exact, while the conservations of N, L_μ and L_e are only approximate. Hence, very slow transitions such as $p \rightleftarrows e^+ + \pi^o$, $\nu_\mu \rightleftarrows \nu_e$, etc. may be allowed. [The present lifetime limit*** of the proton is $> 2 \times 10^{30}$ years.]

---

* G. Danby, J.-M. Gaillard, K. Goulianos, L. M. Lederman, N. Mistry, M. Schwartz and J. Steinberger, Phys. Rev. Lett. 9, 36 (1962).

** T. D. Lee and C. N. Yang, Phys. Rev. 98, 1501 (1955).

*** K. Landé et al., to be published in the proceedings of the Neutrino '80 conference, Erice.

Problem 10.1. Discuss the experimental foundation of the conservation of N, Q, L_μ and L_e.

Problem 10.2. Prove that in QED, if we define

$$C \equiv \exp\left[i\pi \sum (a^\dagger_{\vec{p},s} a_{\vec{p},s} + a^\dagger_{\vec{p},s} a_{\vec{p},s})\right]$$
$$\cdot \exp\left[\frac{\pi}{2} \sum (b^\dagger_{\vec{p},s} a_{\vec{p},s} - a^\dagger_{\vec{p},s} b_{\vec{p},s})\right] ,$$

where $a_{\vec{p},s}$, $a_{\vec{p},s}$ and $b_{\vec{p},s}$ are respectively the annihilation operators of γ, e^- and e^+ with momentum $\vec{p}$ and helicity s, then C satisfies

$$C\, a_{\vec{p},s}\, C^\dagger = - a_{\vec{p},s} ,$$
$$C\, a_{\vec{p},s}\, C^\dagger = b_{\vec{p},s} ,$$
$$C\, b_{\vec{p},s}\, C^\dagger = a_{\vec{p},s} ,$$
$$C^\dagger C = C^2 = 1 ,$$

and therefore C is the charge conjugation operator.

Hint: Define $M_\theta \equiv \exp\left[\theta \sum (b^\dagger_{\vec{p},s} a_{\vec{p},s} - a^\dagger_{\vec{p},s} b_{\vec{p},s})\right]$.

Through differentiation, show that

$$M_\theta\, a_{\vec{p},s}\, M_\theta^\dagger = \cos\theta\ a_{\vec{p},s} + \sin\theta\ b_{\vec{p},s} ,$$
$$M_\theta\, b_{\vec{p},s}\, M_\theta^\dagger = -\sin\theta\ a_{\vec{p},s} + \cos\theta\ b_{\vec{p},s} .$$

Problem 10.3. Any three-dimensional rotation $r_i \to r'_i = u_{ij} r_j$ can be represented by a two-dimensional rotation $\vec{\theta}$ where the magnitude of $\vec{\theta}$ is the angle of rotation and its direction the axis of rotation (with the convention that when $\vec{\theta} \to 0$, $\vec{r}' = \vec{r} + \vec{\theta} \times \vec{r}$).

(i) Show that, for a spin-$\frac{1}{2}$ field, the rotational operator $\exp(i\vec{J}\cdot\vec{\theta})$, where $\vec{J}$ is the angular momentum operator given by (3.62), satisfies

$$\exp(-i\vec{J}\cdot\vec{\theta})\,\psi(\vec{r},t)\,\exp(i\vec{J}\cdot\vec{\theta}) = e^{i\frac{1}{2}\vec{\sigma}\cdot\vec{\theta}}\,\psi(\vec{r}',t) \quad .$$

(ii) Verify that the QED Hamiltonian in the Coulomb gauge is invariant under the unitary transformation R defined by

$$R\,\psi(\vec{r},t)\,R^{\dagger} = e^{i\frac{1}{2}\vec{\sigma}\cdot\vec{\theta}}\,\psi(\vec{r}',t) \quad ,$$

$$R\,A_i(\vec{r},t)\,R^{\dagger} = u_{ij}\,A_j(\vec{r}',t)$$

and

$$R\,\Pi_i(\vec{r},t)\,R^{\dagger} = u_{ij}\,\Pi_j(\vec{r}',t) \quad .$$

Give explicitly the transformation properties of the annihilation and creation operators in the momentum space under R.

Note that if we write $R = \exp(-i\vec{\mathcal{J}}\cdot\vec{\theta})$, then $\vec{\mathcal{J}}$ is the total angular momentum operator.

Problem 10.4. Consider a Lorentz transformation $x'_\mu = u_{\mu\nu}\,x_\nu$ in the (z,t) plane where $u_{11} = u_{22} = 1$, $u_{33} = (1-v^2)^{-\frac{1}{2}}$, etc. In the interaction representation, the corresponding transformation on the field operators is given by

$$L\,\psi(x)\,L^{-1} = \exp\left(\frac{i}{2}\,\theta\,\gamma_3\,\gamma_4\right)\psi(x')$$

and

$$L\,A_\mu(x)\,L^{-1} = u_{\mu\nu}\,A_\nu(x')$$

where

$$\cosh\theta = (1-v^2)^{-\frac{1}{2}} \quad .$$

Show that:

(i) L is unitary, although the 4×4 matrix $\exp\left(\frac{i}{2}\,\theta\,\gamma_3\,\gamma_4\right)$ is not,

(ii) the interaction Lagrangian density $e j_\mu A_\mu$ is explicitly Lorentz-invariant, where $j_\mu = i\psi^\dagger \gamma_4 \gamma_\mu \psi$, and

(iii) the operator $\psi^c(x)$ transforms in the same way as $\psi(x)$ under an arbitrary Lorentz transformation where

$$\psi_a^c(x) \equiv (\gamma_2)_{a\beta} \psi_\beta^\dagger(x) \quad .$$

Remark: The use of the interaction representation is by no means necessary; it is only for convenience.

Chapter 11

ISOTOPIC SPIN AND G PARITY

11.1 Isospin

The concept of isospin was introduced in the early '30's by Heisenberg to describe the approximate charge-independent nature of the strong interaction between protons and neutrons. From a phenomenological point of view we can adopt the usage of the proton field ψ_p and the neutron field ψ_n. Together they can be represented by a column matrix

$$\psi(x) \equiv \begin{pmatrix} \psi_p(x) \\ \psi_n(x) \end{pmatrix} . \tag{11.1}$$

Each of the fields in turn is a quantized Dirac spinor operator whose indices are suppressed here, for clarity. Although, as we shall discuss (in Part II B on interactions), neither the proton nor the neutron is elementary; so far as their symmetry properties are concerned it suffices to represent them by phenomenological field operators.

1. U_2 symmetry

Let us consider the following linear transformation

$$\psi \rightarrow \psi' = u\psi \tag{11.2}$$

where u is a 2×2 matrix. In order to preserve the anti-commutation relation

$$\{\psi_i(\vec{r}, t),\ \psi_j(\vec{r}', t)\} = \delta_{ij}\,\delta^3(\vec{r} - \vec{r}')$$

where i and j can be p or n, the matrix u must be unitary. The group spanned by all such u's is the U_2 group. As we shall see shortly, corresponding to each u there exists a unitary operator U in the Hilbert space such that

$$U\,\psi(x)\,U^\dagger = u\,\psi(x) \quad . \tag{11.3}$$

We may express ψ_p and ψ_n in terms of the Fourier expansion (3.32):

$$\psi_p = \frac{1}{\sqrt{\Omega}} \sum_{\vec{k},s} \left(a_{\vec{k},s}(p)\, u_{\vec{k},s}\, e^{i\vec{k}\cdot\vec{r}} + b^\dagger_{\vec{k},s}(p)\, v_{\vec{k},s}\, e^{-i\vec{k}\cdot\vec{r}}\right),$$

$$\psi_n = \frac{1}{\sqrt{\Omega}} \sum_{\vec{k},s} \left(a_{\vec{k},s}(n)\, u_{\vec{k},s}\, e^{i\vec{k}\cdot\vec{r}} + b^\dagger_{\vec{k},s}(n)\, v_{\vec{k},s}\, e^{-i\vec{k}\cdot\vec{r}}\right). \tag{11.4}$$

Substituting these expressions into (11.3), we obtain

$$U \begin{pmatrix} a_{\vec{k},s}(p) \\ a_{\vec{k},s}(n) \end{pmatrix} U^\dagger = u \begin{pmatrix} a_{\vec{k},s}(p) \\ a_{\vec{k},s}(n) \end{pmatrix} \tag{11.5}$$

and

$$U \begin{pmatrix} b^\dagger_{\vec{k},s}(p) \\ b^\dagger_{\vec{k},s}(n) \end{pmatrix} U^\dagger = u \begin{pmatrix} b^\dagger_{\vec{k},s}(p) \\ b^\dagger_{\vec{k},s}(n) \end{pmatrix} \quad . \tag{11.6}$$

By writing u explicitly as

$$u = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \tag{11.7}$$

and by calling $U a_{\vec{k},s}(p)\, U^\dagger$ and $U a_{\vec{k},s}(n)\, U^\dagger$, respectively, $a'_{\vec{k},s}(p)$ and $a'_{\vec{k},s}(n)$, we have from (11.5)

$$a'_{\vec{k},s}(p) = u_{11}\, a_{\vec{k},s}(p) + u_{12}\, a_{\vec{k},s}(n)$$

and

$$a'_{\vec{k},s}(n) = u_{21}\, a_{\vec{k},s}(p) + u_{22}\, a_{\vec{k},s}(n) \quad .$$

Their Hermitian conjugates are

$$\begin{aligned} a'^{\dagger}_{\vec{k},s}(p) &= u_{11}^{*}\, a^{\dagger}_{\vec{k},s}(p) + u_{12}^{*}\, a^{\dagger}_{\vec{k},s}(n) \quad , \\ a'^{\dagger}_{\vec{k},s}(n) &= u_{21}^{*}\, a^{\dagger}_{\vec{k},s}(p) + u_{22}^{*}\, a^{\dagger}_{\vec{k},s}(n) \end{aligned} \tag{11.8}$$

or simply

$$U \begin{pmatrix} a^{\dagger}_{\vec{k},s}(p) \\ a^{\dagger}_{\vec{k},s}(n) \end{pmatrix} U^{\dagger} = u^{*} \begin{pmatrix} a^{\dagger}_{\vec{k},s}(p) \\ a^{\dagger}_{\vec{k},s}(n) \end{pmatrix} \quad . \tag{11.9}$$

Likewise, by calling $U\, b^{\dagger}_{\vec{k},s}(p)\, U^{\dagger}$ and $U\, b^{\dagger}_{\vec{k},s}(n)\, U^{\dagger}$, $b'^{\dagger}_{\vec{k},s}(p)$ and $b'^{\dagger}_{\vec{k},s}(n)$, we have

$$\begin{aligned} b'^{\dagger}_{\vec{k},s}(p) &= u_{11}\, b^{\dagger}_{\vec{k},s}(p) + u_{12}\, b^{\dagger}_{\vec{k},s}(n) \quad , \\ b'^{\dagger}_{\vec{k},s}(n) &= u_{21}\, b^{\dagger}_{\vec{k},s}(p) + u_{22}\, b^{\dagger}_{\vec{k},s}(n) \quad . \end{aligned} \tag{11.10}$$

We shall now show that for each u there indeed exists a unitary operator U in the Hilbert space which satisfies (11.3). Let us choose the basis vectors in the Hilbert space to be

$$|0>,\ a^{\dagger}_{\vec{k},s}(p)\,|0>,\ b^{\dagger}_{\vec{k},s}(p)\,|0>,\ a^{\dagger}_{\vec{k},s}(n)\,|0>,\ b^{\dagger}_{\vec{k},s}(n)\,|0>, \cdots$$

$$\begin{aligned} |N_p, N_{\bar{p}}, N_n, N_{\bar{n}}> = \prod_{i=1}^{N_p} \prod_{j=1}^{N_{\bar{p}}} \prod_{\ell=1}^{N_n} \prod_{m=1}^{N_{\bar{n}}} \\ \cdot a^{\dagger}_{\vec{k}_i,s_i}(p)\, b^{\dagger}_{\vec{k}_j,s_j}(p)\, a^{\dagger}_{\vec{k}_\ell,s_\ell}(n)\, b^{\dagger}_{\vec{k}_m,s_m}(n)\,|0>, \end{aligned} \tag{11.11}$$

etc., where $|0>$ satisfies

$$a_{\vec{k},s}(p)\,|0> = a_{\vec{k},s}(n)\,|0> = b_{\vec{k},s}(p)\,|0> = b_{\vec{k},s}(n)\,|0> = 0\,. \tag{11.12}$$

The mapping

$$|0> \to |0>\,, \quad a^{\dagger}_{\vec{k},s}(p)\,|0> \to a'^{\dagger}_{\vec{k},s}(p)\,|0>\,,$$

$$b^{\dagger}_{\vec{k},s}(p)\,|\,0\rangle \to b'^{\dagger}_{\vec{k},s}(p)\,|\,0\rangle, \quad a^{\dagger}_{\vec{k},s}(n)\,|\,0\rangle \to a'^{\dagger}_{\vec{k},s}(n)\,|\,0\rangle \ ,$$

$$b^{\dagger}_{\vec{k},s}(n)\,|\,0\rangle \to b'^{\dagger}_{\vec{k},s}(n)\,|\,0\rangle, \ \cdots ,$$

$$|\,N_p, N_{\bar{p}}, N_n, N_{\bar{n}}\rangle \to \prod_{i=1}^{N_p}\prod_{j=1}^{N_{\bar{p}}}\prod_{\ell=1}^{N_n}\prod_{m=1}^{N_{\bar{n}}} \cdot a'^{\dagger}_{\vec{k}_i,s_i}(p)\, b'^{\dagger}_{\vec{k}_j,s_j}(\bar{p})\, a'^{\dagger}_{\vec{k}_\ell,s_\ell}(n)\, b'^{\dagger}_{\vec{k}_m,s_m}(\bar{n})\,|\,0\rangle \qquad (11.13)$$

etc., clearly preserves the orthonormality relations between these vectors, and therefore it is a unitary transformation. Furthermore, the unitary transformation matrix satisfies (11.6) and (11.9) and consequently also (11.3). This establishes that U exists and is unitary. In addition, on account of the first expression in (11.13),

$$U\,|\,0\rangle = |\,0\rangle \ . \qquad (11.14)$$

2. Isospin transformations

Let us separate from U the phase factor (10.118), related to the baryon conservation. We write

$$U = e^{iN\theta}\,S \ , \qquad (11.15)$$

with the corresponding u as

$$u = e^{-i\theta}\,s \ ;$$

therefore (11.3) can be converted to

$$S\,\psi(x)\,S^{\dagger} = s\,\psi(x) \ , \qquad (11.16)$$

where θ is chosen such that

$$\det|\,s\,| = 1 \ . \qquad (11.17)$$

The group $\{s\}$, spanned by all such 2×2 unitary matrices with unit determinant, is the isospin-SU_2 group; on account of (11.16), so is the group $\{S\}$. The properties of $\{s\}$ are exactly the same

as those of the usual three-dimensional rotational group (with spinors), except for the replacement of the ordinary space by the isospin space.

The strong interaction Hamiltonian H_{st} is assumed to be invariant under this SU_2 transformation

$$S\, H_{st}\, S^{\dagger} = H_{st} \quad , \tag{11.18}$$

which will be referred to as the isospin transformation. Since p and n have different electromagnetic interactions, clearly the electromagnetic interaction violates the isospin invariance. We may decompose H_{st} into

$$H_{st} = H_{free} + H_{int} \quad . \tag{11.19}$$

Both H_{free} and H_{int} are isospin-symmetric, with the basis vectors (11.11) as eigenstates of H_{free}. In the approximation that the electromagnetic and weak interactions are neglected, the physical vacuum state $|\, vac >$ as well as the physical nucleon or antinucleon states are all eigenvectors of H_{st} ; on account of (6.58) and (6.60) these states are given by

$$\begin{aligned}
|\, vac > &= U(0,-\infty)\, |\, 0 > \quad , \\
|\, p > &\equiv U(0,-\infty)\, a^{\dagger}_{\vec{k},s}(p)\, |\, 0 > \quad , \\
|\, n > &\equiv U(0,-\infty)\, a^{\dagger}_{\vec{k},s}(n)\, |\, 0 > \quad , \\
|\, \bar{p} > &\equiv U(0,-\infty)\, b^{\dagger}_{\vec{k},s}(p)\, |\, 0 > \quad , \\
|\, \bar{n} > &\equiv U(0,-\infty)\, b^{\dagger}_{\vec{k},s}(n)\, |\, 0 > \quad .
\end{aligned} \tag{11.20}$$

Equations (11.14) and (11.15) imply

$$S\, |\, 0 > = |\, 0 > \quad . \tag{11.21}$$

Because of (11.18), S commutes with $U(0,-\infty)$. Hence

$$S \mid \text{vac} > \; = \; U(0, -\infty)\; S \mid 0 > \; = \; \mid \text{vac} > \quad . \tag{11.22}$$

Likewise, from (11.6), (11.9), (11.15) and (11.16) we have

$$S \begin{pmatrix} a^{\dagger}_{\vec{k},s}(n) \\ a^{\dagger}_{\vec{k},s}(n) \end{pmatrix} S^{\dagger} \; = \; s^{*} \begin{pmatrix} a^{\dagger}_{\vec{k},s}(p) \\ a^{\dagger}_{\vec{k},s}(n) \end{pmatrix}$$

and (11.23)

$$S \begin{pmatrix} b^{\dagger}_{\vec{k},s}(p) \\ b^{\dagger}_{\vec{k},s}(n) \end{pmatrix} S^{\dagger} \; = \; s \begin{pmatrix} b^{\dagger}_{\vec{k},s}(p) \\ b^{\dagger}_{\vec{k},s}(n) \end{pmatrix}$$

which, together with (11.20), gives

$$S \begin{pmatrix} p \\ n \end{pmatrix} \; = \; s^{*} \begin{pmatrix} p \\ n \end{pmatrix} \qquad \text{and} \qquad S \begin{pmatrix} \bar{p} \\ \bar{n} \end{pmatrix} \; = \; s \begin{pmatrix} \bar{p} \\ \bar{n} \end{pmatrix} \tag{11.24}$$

where

$$\begin{pmatrix} p \\ n \end{pmatrix} \; \equiv \; \begin{pmatrix} \mid p > \\ \mid n > \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} \bar{p} \\ \bar{n} \end{pmatrix} \; \equiv \; \begin{pmatrix} \mid \bar{p} > \\ \mid \bar{n} > \end{pmatrix} \quad . \tag{11.25}$$

Let us consider an infinitesimal isospin transformation; i.e.

$$s \; = \; 1 + \tfrac{1}{2} i \vec{\tau} \cdot \vec{\epsilon} \tag{11.26}$$

where the components of $\vec{\tau}$ are the Pauli matrices given by (3.1) and $\vec{\epsilon}$ is an infinitesimal vector. Correspondingly, the Hilbert space transformation S must also be infinitesimally close to the unit matrix; we may therefore write

$$S \; = \; 1 \; - \; i \vec{I} \cdot \vec{\epsilon} \tag{11.27}$$

which can also be regarded as the definition of $\vec{I}$, and will be referred to as the <u>isospin operator</u>. Because s is unitary and the Pauli matrices are all Hermitian,

$$s\, s^{\dagger} \; = \; 1 + \tfrac{1}{2} i \vec{\tau} \cdot (\vec{\epsilon} - \vec{\epsilon}^{\,*}) + O(\epsilon^{2}) \; = \; 1 \quad ,$$

and therefore

$$\vec{\epsilon} = \vec{\epsilon}^{\,*} \quad . \tag{11.28}$$

The unitarity of S then implies

$$S^\dagger S = 1 + i\vec{\epsilon}\cdot(\vec{I}^\dagger - \vec{I}) + O(\epsilon^2) = 1 \tag{11.29}$$

which leads to the Hermiticity of the isospin operator

$$\vec{I} = \vec{I}^{\,\dagger} \quad . \tag{11.30}$$

By substituting (11.27) into (11.18), we find

$$S H_{st} S^\dagger = H_{st} - i\vec{\epsilon}\cdot(\vec{I} H_{st} - H_{st}\vec{I}) + O(\epsilon^2) = H_{st} \ ,$$

which gives

$$[\vec{I}, H_{st}] = 0 \quad . \tag{11.31}$$

A similar substitution of (11.27) into (11.16) results in

$$S\psi S^\dagger = \psi - i\vec{\epsilon}\cdot(\vec{I}\psi - \psi\vec{I}) + O(\epsilon^2) = (1 + \tfrac{1}{2} i\vec{\tau}\cdot\vec{\epsilon})\,\psi \ ,$$

and therefore

$$[\psi(x), \vec{I}] = \tfrac{1}{2}\vec{\tau}\,\psi(x) \quad . \tag{11.32}$$

We shall now show that for a finite isospin transformation

$$s = e^{i\frac{1}{2}\vec{\tau}\cdot\vec{\theta}} \ , \tag{11.33}$$

the corresponding Hilbert space transformation matrix is

$$S = e^{-i\vec{I}\cdot\vec{\theta}} \quad . \tag{11.34}$$

Proof. Let

$$\theta \equiv |\vec{\theta}| \ , \quad \hat{\theta} \equiv \frac{\vec{\theta}}{\theta} \ , \quad I_o \equiv \vec{I}\cdot\hat{\theta} \ , \quad \tau_o \equiv \vec{\tau}\cdot\hat{\theta}$$

and

$$\psi_\theta(x) \equiv e^{-i I_o \theta}\,\psi(x)\, e^{i I_o \theta} \quad . \tag{11.35}$$

The derivative of $\psi_\theta(x)$ is, on account of (11.32) and (11.35),

$$\frac{d}{d\theta}\psi_\theta(x) = -i\, e^{-i I_o \theta}[I_o, \psi(x)]\, e^{i I_o \theta} = \tfrac{1}{2} i\tau_o \psi_\theta(x) \ ,$$

which, upon integration gives $\psi_\theta(x) = e^{i\frac{1}{2}\vec{\tau}\cdot\vec{\theta}}\,\psi(x)$; consequently

$$e^{-i\vec{I}\cdot\vec{\theta}}\,\psi(x)\,e^{i\vec{I}\cdot\vec{\theta}} = e^{i\frac{1}{2}\vec{\tau}\cdot\vec{\theta}}\,\psi(x) \quad . \tag{11.36}$$

That completes the proof.

If the system consists only of nucleons and antinucleons, then by using (3.24) and (11.32) we can easily verify that the isospin operator $\vec{I}$ is given by

$$\vec{I} = \vec{I}_N = \tfrac{1}{2}\int \psi^\dagger(x)\,\vec{\tau}\,\psi(x)\,d^3r \tag{11.37}$$

where, as before, $x_\mu = (\vec{r}, it)$. Because of (11.31), $\vec{I}$ is t-independent and therefore the choice of t in the integrand of (11.37) is arbitrary. If the system contains other particles such as mesons, $\cdots$, then $\vec{I} = \vec{I}_N + \vec{I}_{\text{Mesons}} + \cdots$ where $\vec{I}_N$ is given by (11.37) and $\vec{I}_{\text{Mesons}} + \cdots$ commutes with the nucleon field $\psi(x)$.

Remarks. Since H_{st} is isospin-invariant, it can be diagonalized simultaneously with $\vec{I}^2$ and I_3. Let $I(I+1)$ be the eigenvalue of the operator $\vec{I}^2$. Thus, each hadron carries the quantum numbers I and I_3. For a given I, I_3 can vary from $-I$ to I, making a total of $2I+1$ states. Under isospin rotations, the quantum number I is preserved; however, these $2I+1$ states of different I_3 transform among each other, and therefore they are degenerate with respect to the strong interaction. This mass degeneracy of hadrons is lifted by the electromagnetic and weak interactions; both violate the isospin symmetry.

Exercise. Show that $[I_i, I_j] = i\,\epsilon_{ijk}\,I_k$ where ϵ_{ijk} is given by (3.4).

11.2 G Parity

1. Nucleon-antinucleon system

The proton p and the neutron n form an isospin doublet; e.g. we may regard p as $\uparrow$ and n as $\downarrow$, corresponding to $I_3 = \frac{1}{2}$ and $-\frac{1}{2}$ respectively. Therefore, apart from a phase factor, $\bar{p}$ should behave like $\downarrow$ and $\bar{n}$ like $\uparrow$. As will be shown, the precise relations are: for any transformation (11.24) of the nucleon states

$$S\begin{pmatrix} p \\ n \end{pmatrix} = s^* \begin{pmatrix} p \\ n \end{pmatrix} , \tag{11.38}$$

we have an identical transformation of the antinucleon states

$$S\begin{pmatrix} \bar{n} \\ -\bar{p} \end{pmatrix} = s^* \begin{pmatrix} \bar{n} \\ -\bar{p} \end{pmatrix} . \tag{11.39}$$

Proof. For an arbitrary isospin transformation, we can write, in accordance with (11.33),

$$s = e^{i\frac{1}{2}\vec{\tau}\cdot\vec{\theta}} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(i\tfrac{1}{2}\vec{\tau}\cdot\vec{\theta}\right)^n . \tag{11.40}$$

In view of the fact that for any integer n

$$(\vec{\tau}\cdot\vec{\theta})^{2n} = \theta^{2n} \quad \text{and} \quad (\vec{\tau}\cdot\vec{\theta})^{2n+1} = \theta^{2n}\vec{\tau}\cdot\vec{\theta} ,$$

where $\theta = |\vec{\theta}|$, (11.40) becomes

$$s = \cos\frac{\theta}{2} + i\vec{\tau}\cdot\hat{\theta}\sin\frac{\theta}{2} \tag{11.41}$$

with $\hat{\theta} = \vec{\theta}/\theta$. From (3.1) we see that

$$\tau_2\vec{\tau} = -\vec{\tau}^*\tau_2 ,$$

and therefore

$$\tau_2 s = s^*\tau_2 . \tag{11.42}$$

Multiplying the second equation in (11.24) by τ_2 on the left, we have

$$\tau_2 S\begin{pmatrix} \bar{p} \\ \bar{n} \end{pmatrix} = \tau_2 s \begin{pmatrix} \bar{p} \\ \bar{n} \end{pmatrix} ,$$

which, together with (11.42), gives

$$S\tau_2\begin{pmatrix}\bar{p}\\ \bar{n}\end{pmatrix} = s^*\tau_2\begin{pmatrix}\bar{p}\\ \bar{n}\end{pmatrix} ,$$

and that establishes (11.39).

The same result can also be expressed in terms of the field operators. From (11.16), we have

$$S\begin{pmatrix}\psi_p\\ \psi_n\end{pmatrix}S^\dagger = s\begin{pmatrix}\psi_p\\ \psi_n\end{pmatrix} , \tag{11.43}$$

which leads to

$$S\begin{pmatrix}\psi_p^\dagger\\ \psi_n^\dagger\end{pmatrix}S^\dagger = s^*\begin{pmatrix}\psi_p^\dagger\\ \psi_n^\dagger\end{pmatrix} .$$

Multiplying the above equation by τ_2 on the left and using (11.42), we derive

$$S\begin{pmatrix}\psi_n^c\\ -\psi_p^c\end{pmatrix}S^\dagger = s\begin{pmatrix}\psi_n^c\\ -\psi_p^c\end{pmatrix} \tag{11.44}$$

where

$$(\psi_p^c)_a = (\gamma_2)_{a\beta}(\psi_p^\dagger)_\beta \quad \text{and} \quad (\psi_n^c)_a = (\gamma_2)_{a\beta}(\psi_n^\dagger)_\beta$$

Remarks. If we adopt the notation that $\uparrow$ and $\downarrow$ stand for the $I_3 = \frac{1}{2}$ and $-\frac{1}{2}$ states of an isospin doublet, then from (11.38)-(11.39) the nucleon doublet can be written as

$$p = \uparrow , \qquad n = \downarrow , \tag{11.45}$$

while the antinucleon doublet is

$$\bar{n} = \uparrow \text{ and } \qquad -\bar{p} = \downarrow . \tag{11.46}$$

2. The quantum number G

Whenever several conservation laws operate for the same system it is often possible to obtain new quantum numbers and new

selection rules, as will be illustrated by the interplay between the isospin symmetry and the C invariance of the strong interaction. Under the particle-antiparticle conjugation C we have

$$\begin{pmatrix} p \\ n \end{pmatrix} \rightarrow \begin{pmatrix} \bar{p} \\ \bar{n} \end{pmatrix} .$$

Take any axis $\perp$ to the third axis in the isospin space. A 180° rotation along that axis would transform, apart from some important phase factors,

$$\uparrow \rightarrow \downarrow \quad \text{and} \quad \downarrow \rightarrow \uparrow$$

where, as before, $\uparrow$ can be p or $\bar{n}$ and $\downarrow$ can be n or $\bar{p}$. Now if we can pick the "right" axis and be in full control of the phase factors so that the above transformation becomes precisely

$$\begin{pmatrix} \bar{p} \\ \bar{n} \end{pmatrix} \rightarrow \begin{pmatrix} \bar{n} \\ -\bar{p} \end{pmatrix} ,$$

then, because of (11.38)-(11.39), the chain action of C together with this appropriate 180° isospin rotation would leave the isospin properties of the states invariant, which in turn can lead to a new quantum number, as we shall show.

It is clear that we must be careful about the phase factors. Let us adopt the convention that the nucleon fields ψ_p and ψ_n transform in the same way under C ; i.e.,

$$C \psi_p C^\dagger = \psi_p^c \qquad \text{and} \qquad C \psi_n C^\dagger = \psi_n^c \tag{11.47}$$

or, in terms of the state vector

$$C \mid p > = \mid \bar{p} > \quad \text{and} \quad C \mid n > = \mid \bar{n} > .$$

We define the G parity operator to be

$$G = C \cdot e^{i \pi I_2} , \tag{11.48}$$

where, on account of (11.36),

$$e^{i\pi I_2}\,\psi\,e^{-i\pi I_2} = s\psi \quad , \tag{11.49}$$

with

$$s = e^{-i\frac{1}{2}\pi\tau_2} = -i\tau_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = s^* \quad . \tag{11.50}$$

From (11.38)-(11.39), we also have

$$e^{i\pi I_2}\begin{pmatrix} p \\ n \end{pmatrix} = \begin{pmatrix} -n \\ p \end{pmatrix}$$

and

$$e^{i\pi I_2}\begin{pmatrix} \bar{n} \\ -\bar{p} \end{pmatrix} = \begin{pmatrix} \bar{p} \\ \bar{n} \end{pmatrix} \quad . \tag{11.51}$$

Theorem. $$[G, \vec{I}] = 0 \quad . \tag{11.52}$$

Proof. Let us first consider the single nucleon or antinucleon state. By applying C onto the lefthand sides of (11.51) and (11.52), we find

$$G\begin{pmatrix} p \\ n \end{pmatrix} = -\begin{pmatrix} \bar{n} \\ -\bar{p} \end{pmatrix}$$

and

$$G\begin{pmatrix} \bar{n} \\ -\bar{p} \end{pmatrix} = \begin{pmatrix} p \\ n \end{pmatrix} \quad . \tag{11.53}$$

Thus, for any isospin transformation S, we may combine (11.38)-(11.39) with (11.53), and derive

$$SG\begin{pmatrix} p \\ n \end{pmatrix} = -s^*\begin{pmatrix} \bar{n} \\ -\bar{p} \end{pmatrix} = GS\begin{pmatrix} p \\ n \end{pmatrix}$$

and

$$SG\begin{pmatrix} \bar{n} \\ -\bar{p} \end{pmatrix} = s^*\begin{pmatrix} p \\ n \end{pmatrix} = GS\begin{pmatrix} \bar{n} \\ -\bar{p} \end{pmatrix} \quad . \tag{11.54}$$

The same considerations can also be applied to the field operators. From (11.43)-(11.44) and (11.48)-(11.50), it follows that the equivalent form of (11.53) is

$$G\begin{pmatrix} \psi_p \\ \psi_n \end{pmatrix}G^\dagger = -\begin{pmatrix} \psi_n^c \\ -\psi_p^c \end{pmatrix}$$

and

$$G\begin{pmatrix} \psi_n^c \\ -\psi_p^c \end{pmatrix}G^\dagger = \begin{pmatrix} \psi_p \\ \psi_n \end{pmatrix} \quad ; \tag{11.55}$$

and that of (11.54) is

$$SG\begin{pmatrix}\psi_p\\ \psi_n\end{pmatrix}G^\dagger S^\dagger = GS\begin{pmatrix}\psi_p\\ \psi_n\end{pmatrix}S^\dagger G^\dagger$$

and

$$SG\begin{pmatrix}\psi_n^c\\ -\psi_p^c\end{pmatrix}G^\dagger S^\dagger = GS\begin{pmatrix}\psi_n^c\\ -\psi_p^c\end{pmatrix}S^\dagger G^\dagger \quad . \tag{11.56}$$

Next, we consider the arbitrary multi-nucleon antinucleon state $|\ > = |N_p, N_{\bar{p}}, N_n, N_{\bar{n}}>$ appearing in (11.11); from (11.56) we see that

$$SG\,|\ > = GS\,|\ > \quad .$$

Consequently, if the system consists only of nucleons and antinucleons, then

$$[S, G] = 0 \quad . \tag{11.57}$$

To extend the system to include hadrons other than nucleons and antinucleons, we observe that starting from the initial state which consists only of nucleons and antinucleons (as in any realistic high-energy experiment), it is possible to reach all known hadrons through the strong interaction. Because the strong interaction is assumed to be isospin- and C-invariant, the fact that (11.57) is valid for the initial state means that it must also be valid for the final state. Since (11.57) holds for any $S = e^{i\vec{I}\cdot\vec{\theta}}$, the theorem is established.

From (11.53) or (11.55), we see that

$$G^2\begin{pmatrix}p\\ n\end{pmatrix} = -\begin{pmatrix}p\\ n\end{pmatrix}, \qquad G^2\begin{pmatrix}\bar{n}\\ -\bar{p}\end{pmatrix} = -\begin{pmatrix}\bar{n}\\ -\bar{p}\end{pmatrix}$$

and their generalization

$$G^2\,|\ > = \begin{cases} -\,|\ > & \text{if the state is of } I = \frac{1}{2}, \frac{3}{2}, \cdots \\ \ \ |\ > & \text{if the state is of } I = 0, 1, \cdots . \end{cases}$$

Thus, we can write

$$G^2 = (-1)^{2I} \quad . \tag{11.58}$$

Remarks. Under C, the third component of isospin changes sign; e.g., p is of $I_3 = \frac{1}{2}$, while $\bar{p}$ is of $I_3 = -\frac{1}{2}$. Consequently, C does not commute with $\vec{I}$. From C and $\exp(i\pi I_2)$, we derive G. The importance of G lies in its commutativity with $\vec{I}$. Since H_{st} is both isospin- and C-invariant, it is also G-invariant. This enables us to consider states which are eigenvectors of H_{st}, I, I_3 and G, thereby resulting in some rather useful selection rules, as will be discussed in the next section.

11.3 Applications to Mesons and Baryons

In this section we again consider only the strong interaction, and ignore the isospin-violating effects of the electromagnetic and weak interactions. Our discussions will be mainly on the reasoning that determined the quantum numbers of various hadrons. For convenience, we shall sometimes refer to the three components I_1, I_2 and I_3 of the isospin vector also as I_x, I_y and I_z respectively.

1. Pion

Since the virtual transition

$$\pi \rightleftarrows N\bar{N} \tag{11.59}$$

with $N = p$ or n, can occur via the strong interaction, the pion-isospin I_π must be the same as that of $N\bar{N}$; i.e., 0 or 1. The fact that $\pi^\pm$ and π^0 are of approximately the same mass, making a total of $2I + 1 = 3$-fold degeneracy, means that

$$I_\pi = 1 \quad . \tag{11.60}$$

Transition (11.59) then leads to $\pi^+ \rightleftarrows p\bar{n}$ having $I_z = 1$; likewise,

π^o has $I_z = 0$, and π^- $I_z = -1$. Under an isospin rotation, the three pion states transform like a vector $\vec{r} = (x, y, z)$ in the isospin space with

$$| \pi^{\pm} > \sim \mp 2^{-\frac{1}{2}}(x \pm iy) \quad \text{and} \quad | \pi^o > \sim z \ , \qquad (11.61)$$

where the sign convention is chosen in accordance with the spherical harmonics $Y_{1,\pm 1}$ and $Y_{1,0}$ given by (1.40). A 180^o rotation along the y-axis in the isospin space changes z to $-z$, and therefore

$$e^{i\pi I_y} | \pi^o > = - | \pi^o > \ . \qquad (11.62)$$

From (10.76), $\pi^o \to 2\gamma$, we know that

$$C | \pi^o > = | \pi^o > \ . \qquad (11.63)$$

Thus under $G = Ce^{i\pi I_y}$, the state $| \pi^o >$ must change sign; i.e.,

$$G | \pi^o > = - | \pi^o > \ . \qquad (11.64)$$

Because $| \pi^o >$ can be transformed into an arbitrary coherent mixture of the charged pion states through isospin rotations, the commutation relation $[G, \vec{I}] = 0$ implies that the G-quantum number of $\pi^{\pm}$ must be the same as that of π^o. From (11.64), it follows that

$$G | \pi > = - | \pi > \qquad (11.65)$$

where π can be either $\pi^{\pm}$ or π^o; thereby we derive the selection rule that under the strong interaction

$$\text{even number of pions} \not\leftrightarrow \text{odd number of pions}, \qquad (11.66)$$

which is valid independently of the pion charge.

The assignment (11.63) relies on the 2γ decay of π^o. Is it possible to determine the particle-antiparticle conjugation of π^o without relying on the electromagnetic interaction? The answer is "yes", as will be shown by the following exercise.

Exercise. Assuming that the pion is a pseudoscalar, the strong-interaction transition $\pi \leftrightarrows N\bar{N}$ can be described by the phenomenological Lagrangian density

$$\mathcal{L}_{int} = i g_{\pi N} \psi^{\dagger} \gamma_4 \gamma_5 \vec{\tau} \psi \cdot \vec{\phi} \tag{11.67}$$

where $g_{\pi N}$ is the π-nucleon coupling constant, ψ the nucleon field given by (11.1) and $\vec{\phi}$ the isovector pion field.

(i) Show that, because of (11.47), in order for $\mathcal{L}_{int}$ to be C-invariant, the z-component of $\vec{\phi}$, which represents π^o, must obey

$$\phi_z = C \phi_z C^{\dagger} \quad .$$

Hence, π^o is of $C = +1$, which is now determined by the strong interaction alone.

(ii) Work out separately the transformations of $\vec{\phi}$ under C, $\exp(i\pi I_y)$ and G.

2. Vector mesons

Let us consider the $e^+ e^-$ colliding-beam experiment:

$$e^+ + e^- \rightarrow \text{virtual } \gamma \rightarrow \begin{cases} \rho^o & (770) \\ \omega^o & (780) \\ \phi^o & (1020) \\ J/\psi & (3100) \\ \psi' & (3700) \\ \Upsilon & (9500) \end{cases} \tag{11.68}$$

where the numbers refer to the masses of the vector mesons in units of MeV.

Because $\rho^{\pm}$ and ρ^o are of approximately the same mass, the ρ-mesons form an isospin triplet; therefore, the isospin of ρ is

$$I_\rho = 1 \quad .$$

All the other mesons in the final states of (11.68), such as ω^o, ϕ^o, J/ψ, ⋯ do not have charged states of approximately the same masses. Consequently, they are all isosinglets: $I_\omega = I_\phi = \cdots = 0$.

Next, we want to show that in reaction (11.68) all the final mesons are of

$$C = -1 \quad , \quad P = -1 \quad \text{and} \quad \text{spin} = 1 \quad . \tag{11.69}$$

Proof. Reaction (11.68) proceeds via the intermediate state of a virtual γ, whose transformation properties are the same as those of the photon field A_μ. Since $C A_\mu C^\dagger = -A_\mu$, a virtual γ must be of $C = -1$, so also must the final meson. However, the spin-parity of a virtual γ can be $1-$ or $0+$, since the space-component of A_μ is a vector while its time-component is a scalar. [For example, the well-known Coulomb excitations in atomic and nuclear physics are all via the time-like component of A_μ.] From this we conclude that the final meson can only be either $1-$ or $0+$. Next we shall show that $0+$ is impossible.

From the Feynman diagram in Fig. 11.1, we see that the amplitude of (11.68) is proportional to the following matrix element of the hadronic electromagnetic current operator $J_\mu(x)$:

$$\langle \text{meson} \mid J_\mu(x) \mid \text{vac} \rangle \tag{11.70}$$

where $J_\mu(x)$ satisfies the current conservation law

$$\frac{\partial J_\mu}{\partial x_\mu} = 0 \quad . \tag{11.71}$$

Let q_μ be the 4-momentum of the final meson and P_μ the 4-momentum operator of the system, with $P_4 = i H$ where H is the Hamiltonian operator. The relativistic generalization of Heisenberg's

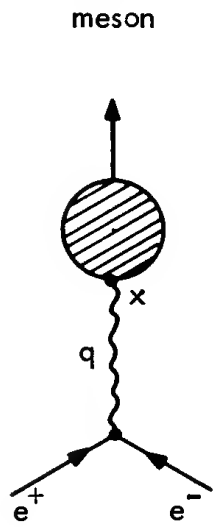


Fig. 11.1. Diagram for $e^+ + e^- \rightarrow$ virtual $\gamma(q) \rightarrow$ meson.

equation (1.9) is

$$[P_\mu, O(x)] = i \frac{\partial O}{\partial x_\mu} \tag{11.72}$$

for any local operator $O(x)$. By taking the matrix element of (11.72) between $|\ \text{vac}\ >$ and $<\ \text{meson}\ |$, we find

$$<\text{meson}\ |\ P_\mu O(x) - O(x)\ P_\mu\ |\ \text{vac}> = i \frac{\partial}{\partial x_\mu} <\text{meson}\ |\ O(x)\ |\ \text{vac}>. \tag{11.73}$$

Since

$$P_\mu\ |\ \text{vac}> = 0 \quad \text{and} \quad P_\mu\ |\ \text{meson}> = q_\mu\ |\ \text{meson}>\ ,$$

(11.73) becomes

$$\frac{\partial}{\partial x_\mu} <\text{meson}\ |\ O(x)\ |\ \text{vac}> = -i\, q_\mu <\text{meson}\ |\ O(x)\ |\ \text{vac}>\ ,$$

which relates the matrix element of $O(x)$ to its value at $x = 0$:

$$<\text{meson}\ |\ O(x)\ |\ \text{vac}> = e^{-iq_\mu x_\mu} <\text{meson}\ |\ O(0)\ |\ \text{vac}>\ . \tag{11.74}$$

Now, if the final meson is of spin-0, then since the matrix

element $< \text{meson} \mid J_\mu(0) \mid \text{vac} >$ depends only on a single 4-vector q_μ and because it is itself a 4-vector, we can write

$$< \text{meson} \mid J_\mu(0) \mid \text{vac} > = c q_\mu$$

where c is a constant. By using (11.74), we find

$$< \text{meson} \mid J_\mu(x) \mid \text{vac} > = c q_\mu e^{-i q_\mu x_\mu}$$

which, together with (11.71), gives

$$0 = \frac{\partial}{\partial x_\mu} < \text{meson} \mid J_\mu(x) \mid \text{vac} > = -i c q_\mu^2 e^{-i q_\mu x_\mu} .$$

Because $-q_\mu^2 = (\text{c.m. energy})^2 \neq 0$, it follows that $c = 0$, i.e.,

$$< \text{meson} \mid J_\mu(x) \mid \text{vac} > = 0 \quad \text{if} \quad \text{spin} = 0 . \tag{11.75}$$

Hence the spin of the final meson cannot be 0, and (11.69) is then established. From C conservation, we conclude that all such mesons obey the selection rule

$$C = -1 \text{ meson} \not\to \text{any number of } \pi^0 . \tag{11.76}$$

Because the π^0's are bosons, we have

$$1^- \text{ meson} \not\to 2\pi^0 . \tag{11.77}$$

Since the ρ^0 meson is of $I = 1$, $I_z = 0$ and $C = -1$, therefore G of ρ^0 is $+1$. The commutation relation (11.52) implies that $\rho^\pm$ is also of $G = +1$. Hence, G conservation requires that independently of the charge

$$\rho \not\to \text{odd number of pions.} \tag{11.78}$$

Likewise, since ω^0, ϕ^0, J/ψ, $\cdots$ are all of $I = 0$, and therefore $G = -1$, they obey the selection rule

$$\omega, \phi, J/\psi, \cdots \not\to \text{even number of pions.} \tag{11.79}$$

As we mentioned earlier, isospin symmetry and G symmetry are

violated by the electromagnetic interaction, and C symmetry by the weak interaction. Hence, the violation amplitude of selection rules (11.78)-(11.79) is of the order of the fine-structure constant, while that of (11.76) is due to the weak interaction and therefore of a much smaller magnitude. So far as we know, selection rule (11.77) is exact.

Exercise. If $|A>$ and $|B>$ are eigenstates of the 4-momentum operator P_μ with eigenvalues a_μ and b_μ, show that the x-dependence of the matrix element $<B|O(x)|A>$ is completely known:

$$<B|O(x)|A> = <B|O(0)|A> e^{i(a_\mu - b_\mu)x_\mu} \quad (11.80)$$

for any local operator $O(x)$.

3. Λ and kaon

There is only one particle which is degenerate with Λ^o, and that is its antiparticle $\bar{\Lambda}^o$. Consequently, the isospin degeneracy $2I_\Lambda + 1$ must be $\leqslant 2$, which means that $I_\Lambda = 0$ or $\frac{1}{2}$.

Consider now the strong reaction

$$\pi^- + p \rightarrow \Lambda^o + K^o \quad . \quad (11.81)$$

If I_Λ were $\frac{1}{2}$, then under an isospin rotation, say $\exp(i\pi I_y)$, Λ^o would transform into its isospin partner $\bar{\Lambda}^o$. Because of the decay

$$\Lambda^o \rightarrow \pi + N \quad , \quad (11.82)$$

Λ^o has a baryon number $= +1$; likewise since

$$\bar{\Lambda}^o \rightarrow \pi + \bar{N} \quad ,$$

the antiparticle $\bar{\Lambda}^o$ has a baryon number $= -1$. When we apply the isospin rotation $\exp(i\pi I_y)$ onto (11.81), we find the lefthand side remains a mixture of pion and nucleon; however, if $I_\Lambda = \frac{1}{2}$, the right-hand side would be a mixture of $\bar{\Lambda}^o$ and kaon, and that would

violate baryon conservation. Hence we conclude that

$$I_\Lambda = 0 \quad . \tag{11.83}$$

Because the isospin of π is 1 and that of the nucleon is 1/2, the total isospin of (11.81) can be either 1/2 or 3/2 which, together with (11.83), gives the kaon isospin $I_K = 1/2$ or $3/2$. Let Q_K, Q_π, Q_Λ and Q_N be the charges of K, π, Λ and N. The charges of π and N are related to their I_z's by

$$Q_\pi = I_z \qquad \text{and} \qquad Q_N = I_z + \tfrac{1}{2} \quad .$$

Since $Q_\Lambda = I_\Lambda = 0$, reaction (11.81) implies that

$$Q_K = Q_{\pi+N} = I_z + \tfrac{1}{2}$$

where I_z refers to that of the kaon. Consequently, if $I_K = 3/2$, the kaon state that has $I_z = 3/2$ must be of charge $3/2 + 1/2 = 2$. Because there is no doubly-charged kaon, I_K cannot be $3/2$. Thus, although there are four degenerate kaon states: K^+, K^-, K^o and $\bar{K}^o$, we conclude that

$$I_K = \tfrac{1}{2} \quad . \tag{11.84}$$

The four kaon states form two isospin doublets

$$\begin{pmatrix} K^+ \\ K^o \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} \bar{K}^o \\ -K^- \end{pmatrix} \quad , \tag{11.85}$$

which transform under the isospin rotation like

$$\begin{pmatrix} p \\ n \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} \bar{n} \\ -\bar{p} \end{pmatrix}$$

respectively.

4. Meson and baryon octets

Reasoning identical to that used in determining the isospin of Λ, (11.83), leads to the conclusion that no baryon can belong to the

same isospin multiplet as its antiparticle. From $\pi^- + p \to \Sigma^- + K^+$ we conclude Σ has a baryon number $+1$, and from the known triplet degeneracy Σ^+, Σ^0 and Σ^- we deduce

$$I_\Sigma = 1 \quad . \tag{11.86}$$

Likewise, the allowed reaction $K^- + p \to \Xi^- + K^+$ implies that the baryon number of Ξ is 1; from the doublet degeneracy, Ξ^0 and Ξ^-, we deduce

$$I_\Xi = \tfrac{1}{2} \quad . \tag{11.87}$$

Similarly, from $K^- + p \to \Lambda^0 + \eta^0$, it follows that the baryon number of η^0 is 0. Since η^0 has no degeneracy, we conclude

$$I_\eta = 0 \quad . \tag{11.88}$$

It is useful to define hypercharge Y and strangeness* S:

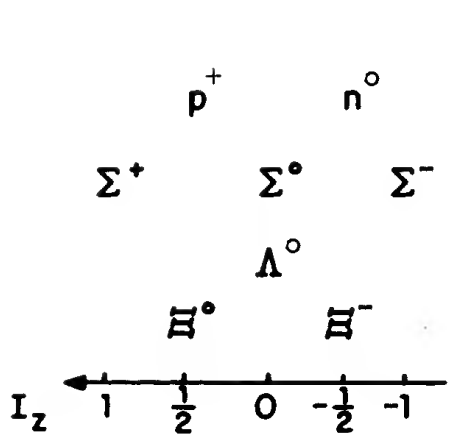
$$\begin{aligned} Y &\equiv 2(Q - I_z) \\ \text{and} \quad S &\equiv Y - N \end{aligned} \tag{11.89}$$

where Q is the electric charge and N the baryon number. Any interaction that conserves Q, I_z and N also conserves Y and S. In Table 11.1 we list the I, I_z, Y and S of the low-lying baryon and meson octets (the word octet refers to the SU_3 representation to be discussed in the next chapter). The baryons are all spin-$\frac{1}{2}$ with even relative parity, while the mesons are all pseudoscalars. We note that p, n and the pions are all of $S = 0$, hence non-strange particles, while Σ, Λ and the kaons have $S = \pm 1$ and therefore are called strange particles.

From Table 11.1 we see that particles with different I_z, but

---

* M. Gell-Mann, Phys.Rev. 93, 933 (1953); T. Nakano and K. Nishijima, Progr.Theoret.Phys. 10, 581 (1953).



	I	Y	S
p^+ n°	$\frac{1}{2}$	1	0
Σ^+ Σ° Σ^-	1	0	-1
Λ°	0	0	-1
Ξ° Ξ^-	$\frac{1}{2}$	-1	-2
K^+ K°	$\frac{1}{2}$	1	1
π^+ π° π^-	1	0	0
η°	0	0	0
$\bar{K}^\circ$ K^-	$\frac{1}{2}$	-1	-1

Table 11.1. Quantum numbers of the pseudoscalar meson and the spin-$\frac{1}{2}$ baryon octets.

the same isospin multiplet, have the same Y, S and N, as expected. Consequently, these quantum numbers commute with $\vec{I}$; i.e.

$$[Y, \vec{I}] = [S, \vec{I}] = [N, \vec{I}] = 0 \quad . \tag{11.90}$$

On the other hand, since $Q = I_z + Y/2$, we have the commutation relations

$$[Q, I_x] = [I_z, I_x] = i I_y \ ,$$

$$[Q, I_y] = [I_z, I_y] = -i I_x \tag{11.91}$$

and

$$[Q, I_z] = 0 \quad .$$

The strong interaction conserves S. As we shall discuss, so does the electromagnetic interaction, but the weak interaction does not. Consequently, a strange particle can decay into non-strange particles only through the weak interaction, and that makes its lifetime long. However, in a collision of non-strange particles such as nucleons and pions, strange particles can be produced copiously through the strong interaction, but only in pairs.* Historically, this explains the apparent paradox between the large production cross sections and long lifetimes of the strange particles.

11.4 Isospin Violation

1. Electromagnetic interaction

As mentioned before, since p and n have different electric charges and I_z, but belong to the same isospin multiplet, this implies that the electromagnetic interaction H_γ cannot be invariant under any isospin rotation that changes I_z; i.e.

$$[H_\gamma, I_x] \neq 0 \qquad \text{and} \qquad [H_\gamma, I_y] \neq 0 \ . \tag{11.92}$$

In transitions such as

$$p \leftrightarrows p + \gamma \qquad \text{and} \qquad n \leftrightarrows n + \gamma \ ,$$

the I_z of the nucleon is unchanged. In other words, these reactions conserve I_z. We will now make the assumption that in all electromagnetic processes I_z is conserved; i.e.

$$[H_\gamma, I_z] = 0 \ . \tag{11.93}$$

The electromagnetic interaction of a particle of charge e can

---

* A. Pais, Phys.Rev. 86, 663 (1952); Y. Nambu, K. Nishijima and Y. Yamaguchi, Progr.Theoret.Phys. 6, 615, 619 (1951).

be obtained most simply by replacing in its free Lagrangian

$$\frac{\partial}{\partial x_\mu} \rightarrow \frac{\partial}{\partial x_\mu} - ieA_\mu \quad . \tag{11.94}$$

The interaction obtained this way automatically satisfies the gauge invariance, and is called the "minimal electromagnetic interaction"; the corresponding electromagnetic current J_μ is called the minimal current. Under the isospin rotation, the minimal current, generated by (11.94), transforms in the same way as the electric-charge operator $Q = -i \int J_4 \, d^3r$. Because $Q = I_z + Y/2$, which consists of an isovector part, I_z, and an isoscalar part, $Y/2$, so does J_μ. We may decompose the minimal electromagnetic current into two terms

$$J_\mu = (J_\mu)_0 + (J_\mu)_1 \tag{11.95}$$

where the subscript 0 or 1 indicates whether the term is of $I = 0$ or 1; both are of $I_z = 0$.

In any transition

$$a \rightarrow b + \gamma$$

where a and b refer to hadron states, we may define the vector of isospin change to be

$$\Delta\vec{I} = \vec{I}_a - \vec{I}_b \quad . \tag{11.96}$$

Hence, under the assumption of minimal electromagnetic interaction and on account of (11.95), we have the following selection rules:

$$\Delta I_z = \Delta Y = \Delta S = 0 \tag{11.97}$$

and

$$|\Delta\vec{I}| = 0 \text{ or } 1 \tag{11.98}$$

where the $\Delta\vec{I} = 0$ transition is due to the isoscalar current $(J_\mu)_0$, while the $|\Delta\vec{I}| = 1$ transition is due to the isovector current $(J_\mu)_1$.

To be sure, gauge invariance alone does not imply minimal interaction. We can always add non-minimal terms to the interaction Lagrangian which depend on $F_{\mu\nu} = \partial A_\nu / \partial x_\mu - \partial A_\mu / \partial x_\nu$, instead of A_μ itself; e.g., an anomalous magnetic-moment term

$$i \psi^\dagger \gamma_4 \gamma_\mu \gamma_\nu \psi' F_{\mu\nu} \quad . \tag{11.99}$$

Such a term is always gauge-invariant; it may or may not be I_z-conserving. If $\psi = \psi_n$ and $\psi' = \psi_\Lambda$, then we could have

$$\Lambda \leftrightarrows n + \gamma \tag{11.100}$$

which violates I_z. Our assumption is that these non-minimal terms do not exist. Experimentally, reaction (11.100) has never been observed, giving support to the assumption of minimal electromagnetic interaction. On the other hand, as a further support of our selection rules,

$$\Sigma^o \leftrightarrows \Lambda + \gamma \tag{11.101}$$

does occur, and is the dominant decay mode of Σ^o. Unlike (11.100), this is an allowed decay since it satisfies $|\Delta \vec{I}| = 1$ and $\Delta I_z = 0$.

Remarks. We note that on account of (11.91) the total charge operator $Q = -i \int J_4 \, d^3r$ commutes with the operator $\vec{I}^2$,

$$[Q, \vec{I}^2] = 0 \quad .$$

Thus, if a matrix element

$$\langle b | Q | a \rangle \neq 0 \quad ,$$

then besides the selection rules (11.96) and (11.97) we must have

$$I_a = I_b$$

where I_a and I_b denote, respectively, the magnitudes of the total isospin of the hadron states a and b. However, the electromagnetic

current operator $J_\mu(x)$ does not commute with $\vec{I}^2$. This is why $\Sigma^0 \rightleftarrows \Lambda + \gamma$ is an allowed transition, in which $I_\Sigma = 1$ but $I_\Lambda = 0$.

2. Weak interaction

Isospin is violated in both the nonleptonic weak processes, such as

$$\Lambda^0 \to p + \pi^- , \quad K^\pm \to \pi^\pm + \pi^0 \tag{11.102}$$

and semileptonic weak processes, such as

$$n \to p + e^- + \bar{\nu}_e , \quad K^\pm \to \pi^0 + e^\pm + \nu_e \ (\text{or } \bar{\nu}_e) . \tag{11.103}$$

In the former we have

$$| \Delta I_z | = \tfrac{1}{2} , \tag{11.104}$$

while in the latter

$$| \Delta I_z | = 1 \ \text{or} \ \tfrac{1}{2} , \tag{11.105}$$

where the change refers only to the ΔI_z of the hadrons. We now make the minimal-violation hypothesis: In all weak processes, the change of $\vec{I}$ of the hadrons satisfies

$$| \Delta \vec{I} | \leqslant 1 . \tag{11.106}$$

Let us see how to apply this rule and why it is called "minimal violation."

(i) Λ-decay

In the decay

$$\Lambda^0 \to \begin{cases} p + \pi^- \\ n + \pi^0 \end{cases} ,$$

the initial state has $I_\Lambda = 0$, while the final-state isospin $I_{N\pi}$ can in principle be $\frac{1}{2}$ or $\frac{3}{2}$. Rule (11.106) states that $I_{N\pi}$ must assume the smaller value, $\frac{1}{2}$; i.e.,

$$| \Delta \vec{I} | = \tfrac{1}{2} . \tag{11.107}$$

As we shall see in the following exercise, such a final state $|\frac{1}{2}>$ can be written as

$$|\tfrac{1}{2}> = \sqrt{\tfrac{1}{3}}\ |n\pi^{o}> - \sqrt{\tfrac{2}{3}}\ |p\pi^{-}> \ . \qquad (11.108)$$

Thus (11.106) leads to the relative decay rates, apart from the electromagnetic correction,

$$\frac{\text{rate } (\Lambda^{o} \to p + \pi^{-})}{\text{rate } (\Lambda^{o} \to n + \pi^{o})} \cong 2 \qquad (11.109)$$

which is consistent with the observed branching ratios of $\Lambda^{o} \to p + \pi^{-}$ and $n + \pi^{o}$ being $(64.2 \pm .5)\%$ and $(35.8 \pm .5)\%$ respectively.

Exercise. In the isospin space we may represent the states p, n and π by

$$|p> = \begin{pmatrix}1\\0\end{pmatrix}, \quad |n> = \begin{pmatrix}0\\1\end{pmatrix}, \quad |\pi^{\pm}> = Y_{1,\pm 1} \quad \text{and} \quad |\pi^{o}> = Y_{1,0} \qquad (11.110)$$

where $Y_{\ell,m}$ refers to the spherical harmonics given by (1.40). The state (11.108) is of $I = \frac{1}{2}$ and $I_z = -\frac{1}{2}$. Show that in the representation of (11.110), it can be written as

$$-\vec{\tau} \cdot \hat{r} \begin{pmatrix}0\\1\end{pmatrix} \qquad (11.111)$$

where $\hat{r} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ in the isospin space. By using the above expression, verify the Clebsch-Gordon coefficients $\sqrt{\frac{2}{3}}$ and $\sqrt{\frac{1}{3}}$ in (11.108).

(ii) $K_{\pi 2}$-decay

In $K^{\pm} \to \pi^{\pm} + \pi^{o}$, K^{o} (or $\bar{K}^{o}$) $\to \pi^{+} + \pi^{-}$ or $2\pi^{o}$, since the spin of K is 0, the final two pions must be in an s-state. Bose statistics then requires that the 2π isospin $I_{2\pi} \neq 1$. Thus, $I_{2\pi}$ can

be 0 or 2. Rule (11.106) states that $I_{2\pi}$ must assume the smaller value

$$I_{2\pi} = 0 \quad .$$

Because $I_K = \frac{1}{2}$, we have again

$$| \Delta \vec{I} | = \tfrac{1}{2} \quad . \tag{11.112}$$

As in the above exercise, we may represent each of the two final pions by a unit vector, $\hat{r}_a$ and $\hat{r}_b$ in the isospin space, with the subscripts referring to, say, their different momenta. The $I_{2\pi} = 0$ state can then be written as being proportional to the scalar product $\hat{r}_a \cdot \hat{r}_b$; i.e.

$$| \pi_a^+ \pi_b^- > + | \pi_a^- \pi_b^+ > + | \pi_a^o \pi_b^o > \quad .$$

Hence we find, apart from the electromagnetic correction,

$$\frac{\text{rate } (K^o \to \pi^+ + \pi^-)}{\text{rate } (K^o \to 2\pi^o)} \cong 2 \tag{11.113}$$

and the same ratio if K^o is replaced by $\bar{K}^o$. Since $\pi^\pm \pi^o$ has $I_z = \pm 1$; it cannot be in an $I_{2\pi} = 0$ state and rule (11.106) implies

$$K^\pm \not\to \pi^\pm + \pi^o \quad . \tag{11.114}$$

The experimental results are

$$\text{branching ratio } (K_S^o \to \pi^+ + \pi^-) = 68.61 \pm .24\% \quad ,$$

$$\text{branching ratio } (K_S^o \to 2\pi^o) = 31.39 \pm .24\%$$

and

$$\frac{\text{rate } (K^\pm \to \pi^\pm + \pi^o)}{\text{rate } (K_S^o \to 2\pi)} = (1.53 \pm 0.07) \times 10^{-3}$$

in agreement with (11.113) and (11.114). [The relation between K^o, $\bar{K}^o$ and $K_S^o \cong (K^o + \bar{K}^o)/\sqrt{2}$ will be discussed in Chapter 15.]

(iii) $K_{\ell 3}$ decay

In the decays

$$K^{\pm} \to \pi^{o} + \ell^{\pm} + \nu_{\ell} \quad (\text{or } \bar{\nu}_{\ell})$$

and

$$K^{o} \ (\text{or } \bar{K}^{o}) \to \pi^{\mp} + \ell^{\pm} + \nu_{\ell} \ (\text{or } \bar{\nu}_{\ell}) \ ,$$

where $\ell = e$ or μ, the initial isospin is $I_K = |\vec{I}_K| = \frac{1}{2}$ while the final isospin is $I_\pi = |\vec{I}_\pi| = 1$. By using the notation of (11.96), we have, according to the usual vector-addition rule of angular momenta,

$$|\Delta\vec{I}| = |\vec{I}_\pi - \vec{I}_K| = \tfrac{1}{2} \text{ or } \tfrac{3}{2} \ . \qquad (11.115)$$

[Note that $|\Delta\vec{I}| \neq \Delta|\vec{I}|$, since $\Delta|\vec{I}| = |\vec{I}_\pi| - |\vec{I}_K| = 1 - \frac{1}{2} = \frac{1}{2}$ always.]

Rule (11.106) now states that between the two possible values of (11.115), we must have, as in (11.107) and (11.112),

$$|\Delta\vec{I}| = \tfrac{1}{2} \qquad (11.116)$$

only. Since the I_z of K^o and π^+ are respectively $-\frac{1}{2}$ and $+1$, the above selection rule implies

$$K^{o} \not\to \pi^{+} + \ell^{-} + \bar{\nu}_{\ell} \ . \qquad (11.117)$$

Likewise,

$$\bar{K}^{o} \not\to \pi^{-} + \ell^{+} + \nu_{\ell} \ . \qquad (11.118)$$

By using (11.116) and by following arguments similar to those that led from (11.110)-(11.111) in the exercise to (11.108), we can derive

$$\frac{\text{rate } (K^{o} \to \pi^{-} + \ell^{+} + \nu_{\ell})}{\text{rate } (K^{+} \to \pi^{o} + \ell^{+} + \nu_{\ell})} \cong 2$$

and (11.119)

$$\frac{\text{rate } (\bar{K}^0 \to \pi^+ + \ell^- + \bar{\nu}_\ell)}{\text{rate } (K^- \to \pi^0 + \ell^- + \bar{\nu}_\ell)} \cong 2 \quad .$$

In terms of $K_L^0 \cong (K^0 - \bar{K}^0)/\sqrt{2}$, which will be discussed in Chapter 15, these results become

$$\frac{\sum \text{ rate } (K_L^0 \to \pi^\pm + \ell^\mp + \nu_\ell \text{ (or } \bar{\nu}_\ell))}{\text{rate } (K^\pm \to \pi^0 + \ell^\pm + \nu_\ell \text{ (or } \bar{\nu}_\ell))} \cong 2 \qquad (11.120)$$

where $\sum$ denotes the sum over + and −. The experimental values are

$$\sum \text{rate } (K_L^0 \to \pi^\pm + \ell^\mp + \nu_\ell \text{ (or } \bar{\nu}_\ell)) = \begin{cases} 5.21 \pm .10 \times 10^6/\text{sec} \\ \qquad \text{for } \ell = \mu \\ 7.49 \pm .11 \times 10^6/\text{sec} \\ \qquad \text{for } \ell = e \end{cases} \qquad (11.121)$$

and

$$\text{rate } (K^\pm \to \pi^0 + \ell^\pm + \nu_\ell \text{ (or } \bar{\nu}_\ell)) = \begin{cases} 2.58 \pm .07 \times 10^6/\text{sec} \\ \qquad \text{for } \ell = \mu \\ 3.90 \pm .04 \times 10^6/\text{sec} \\ \qquad \text{for } \ell = e, \end{cases} \qquad (11.122)$$

which are in good agreement with (11.120). We note that in the decays of the strange particles, the selection rule (11.106) always reduces to $|\Delta \vec{I}| = \frac{1}{2}$.

Problem 11.1. The pion-nucleon scattering at a laboratory π energy ~ 300 MeV is dominated by the resonance $\Delta(1232)$ which is of isospin $\frac{3}{2}$:

$$\pi + N \to \Delta(1232) \to \pi + N \quad .$$

Show that the ratios of the differential cross sections of

$\pi^+ + p \to \pi^+ + p$, $\pi^- + p \to \pi^0 + n$ and $\pi^- + p \to \pi^- + p$ are 9 : 2 : 1 .

Problem 11.2. Prove the selection rules listed in Tables 11.2 and 11.3 following. Discuss the nature of these rules; to what degree is each of them correct?

State:	$^1S_0$	$^3S_1$	$^1P_1$	$^3P_0$	$^3P_1$	$^3P_2$
Spin parity	0−	1−	1+	0+	1+	2+
G	−	+	+	−	−	−
$\pi^- + \pi^0$	×		×	×	×	×
$2\pi^- + \pi^+$		×	×	×		
$\pi^- + 2\pi^0$		×	×	×		
$2\pi^- + \pi^+ + \pi^0$	×			×	×	×
$\pi^- + 3\pi^0$	×			×	×	×
$3\pi^- + 2\pi^+$		×	×			
$2\pi^- + \pi^+ + 2\pi^0$		×	×			
$\pi^- + 4\pi^0$		×	×			

(× means forbidden due to $\vec{I}$, or C, or other selection rules.)

Table 11.2. Selection rules for $p + \bar{n} \to m\pi$.

State:	$^1S_0$		$^3S_1$		$^1P_1$		$^3P_0$		$^3P_1$		$^3P_2$	
Spin parity	0−		1−		1+		0+		1+		2+	
C	+		−		−		+		+		+	
I	0	1	0	1	0	1	0	1	0	1	0	1
G	+	−	−	+	−	+	+	−	+	−	+	−
$2\pi^o$	×	×	×	×	×	×		×	×	×		×
$\pi^+ + \pi^-$	×	×	×		×	×		×	×	×		×
$3\pi^o$	×		×	×	×	×	×	×	×		×	
$\pi^+ + \pi^- + \pi^o$	×			×		×	×	×	×		×	
$4\pi^o$		×	×	×	×	×		×		×		×
$\pi^+ + \pi^- + 2\pi^o$		×	×		×			×		×		×
$2\pi^+ + 2\pi^-$		×	×		×			×		×		×
$5\pi^o$	×		×	×	×	×	×		×		×	
$\pi^+ + \pi^- + 3\pi^o$	×			×		×	×		×		×	
$2\pi^+ + 2\pi^- + \pi^o$	×			×		×	×		×		×	

(× means forbidden due to $\vec{I}$, or C, or other selection rules.)

Table 11.3. Selection rules for $\bar{p} + p \rightarrow m\pi$ or $\bar{n} + n \rightarrow m\pi$.

These two tables are taken from T. D. Lee and C. N. Yang, Nuovo Cimento 3, 749 (1956). G parity was introduced in that paper.

For earlier and related work, see

K. Nishijima, Prog.Theoret.Phys. 6, 614, 1027 (1951).

A. Pais and R. Jost, Phys.Rev. 87, 871 (1952).

L. Michel, Nuovo Cimento 10, 319 (1953).

D. Amati and B. Vitale, Nuovo Cimento 2, 719 (1955).

Chapter 12

SU_3 SYMMETRY

That the strong interaction may have a much wider internal symmetry than the U_2 group was first considered by Sakata*, who explored the possibility of SU_3 symmetry generated by the unitary transformations between p, n and Λ. However, Sakata's approach encountered serious difficulties, since the Λ-nucleon force turns out to be quite different from the nucleon-nucleon force. Major progress was made by Y. Ohnuki** in 1960 who avoided the dynamical difficulties of the Sakata model; instead he put the emphasis on the kinematics of SU_3. The observed hadrons are regarded as composites of a triplet of "baryon" fields, called X_1, X_2 and X_3 by Ohnuki. These fields have the same quantum numbers as p, n and Λ, but their quanta differ from the physical baryons because of some unspecified dynamical bound-state mechanism. By examining various representations of the SU_3 group, Ohnuki was able to identify the physical pions and kaons as members of an SU_3 octet, thereby predicting a new pseudoscalar meson, which was later discovered and is now called η^o.

---

* S. Sakata, Progr.Theoret.Phys. 16, 686 (1956).

** Y. Ohnuki, Proceedings of the International High-energy Conference, CERN (1960), p. 843.

Soon after, Gell-Mann and Ne'eman* took the decisive step of identifying the physical baryons p, n, Λ, Σ and Ξ, also, as members of an SU_3-octet. At that time the basic triplet was regarded more as a mathematical device for the construction of the octet (called the eightfold way) and the decuplet representations, which can then be directly applied to the observed hadrons. These applications of SU_3-symmetry led to great success in bringing order to the complex problems of spectroscopy, dynamics and decay rates of hadrons.

We now know that all hadrons can be viewed as composites of quarks**. In this chapter, we consider only the three low-lying quark fields: up, down and strange, which are sometimes referred to as different "flavors". Each will be represented by an element of the column matrix

$$q = \begin{pmatrix} q^1 \\ q^2 \\ q^3 \end{pmatrix} \quad . \tag{12.1}$$

The strong interaction is assumed to be approximately invariant under the transformation

$$q \rightarrow vq \tag{12.2}$$

where v is unitary. The detailed form of the strong interaction will be considered later when we discuss quantum chromodynamics in Chapter 18. Here we examine only the consequences of the symmetry assumption.

---

* M. Gell-Mann, Phys.Rev. 125, 1067 (1962); Caltech Report CTSL-20 (1961); Y. Ne'eman, Nucl.Phys. 26, 222 (1961).

** G. Zweig, CERN report (unpublished); M. Gell-Mann, Phys.Lett. 8, 214 (1964).

12.1 Mathematical Preliminary

Just as in (11.17), we can separate out the overall phase of v, and consider the SU_3 group $\{v\}$ spanned by all 3×3 unitary matrices with unit determinant; i.e.,

$$v\, v^\dagger = 1 \qquad \text{and} \qquad \det v = 1 \ . \tag{12.3}$$

For convenience, the matrix v will be written as

$$v = (v_b^{\ a}) = \begin{pmatrix} v_1^{\ 1} & v_2^{\ 1} & v_3^{\ 1} \\ v_1^{\ 2} & v_2^{\ 2} & v_3^{\ 2} \\ v_1^{\ 3} & v_2^{\ 3} & v_3^{\ 3} \end{pmatrix} \ ; \tag{12.4}$$

the matrix elements of its Hermitian conjugate $v^\dagger$ are given by

$$v_a^{\dagger b} = (v_b^{\ a})^* \ . \tag{12.5}$$

Thus, the transformation (12.2) can be written as

$$q^a \rightarrow v_b^{\ a}\, q^b \ . \tag{12.6}$$

1. Tensors

Definition: $T_{a_1 \cdots a_m}^{i_1 \cdots i_n}$ is a component of a tensor T of rank (n, m) if its transformation law is

$$T_{a_1 \cdots a_m}^{i_1 \cdots i_n} \rightarrow v_{j_1}^{\ i_1} \cdots v_{j_n}^{\ i_n}\ v_{a_1}^{\dagger b_1} \cdots v_{a_m}^{\dagger b_m}\ T_{b_1 \cdots b_m}^{j_1 \cdots j_n} \ . \tag{12.7}$$

Throughout this section, unless otherwise indicated, all indices vary from 1 to 3. Thus, the above tensor has 3^{n+m} components.

From this definition, by adding or subtracting the corresponding components of two tensors of the same rank, we form a new tensor also of the same rank. By multiplying the components of a tensor A of rank (n, m) with those of a tensor B of rank (q, p), we

can construct a new tensor of rank $(n+q, m+p)$ whose components are

$$A_{a_1 \cdots a_m}^{i_1 \cdots i_n} \; B_{b_1 \cdots b_p}^{j_i \cdots j_q} \quad . \tag{12.8}$$

Isotropic tensors: *

$$\text{(i)} \quad \delta_a^{\,i} = \begin{cases} 1 & \text{if } i = a \ , \\ 0 & \text{otherwise} \ , \end{cases} \tag{12.9}$$

are components of a tensor of rank $(1, 1)$.

Proof. Because $v\, v^{\dagger} = 1$, under the transformation

$$\delta_a^{\,i} \rightarrow v_j^{\,i}\, v^{\dagger b}_{\,a}\, \delta_b^{\,j} = \delta_a^{\,i} \quad .$$

$$\text{(ii)} \quad \epsilon^{ijk} = \begin{cases} 1 & \text{if } ijk \text{ is an even permutation of } 1,2,3, \\ -1 & \text{if } ijk \text{ is an odd permutation of } 1,2,3, \\ 0 & \text{otherwise,} \end{cases} \tag{12.10}$$

are components of a tensor of rank $(3, 0)$.

Proof. Because $\det v = 1$, we have under the transformation

$$\epsilon^{ijk} \rightarrow v_a^{\,i}\, v_b^{\,j}\, v_c^{\,k}\, \epsilon^{abc} = \epsilon^{ijk} \quad .$$

(iii) Likewise, because $\det v^{\dagger} = 1$, ϵ_{ijk} given by (3.4) are components of a tensor of rank $(0, 3)$.

Contraction: From a tensor T of rank (n, m), we can form new tensors of rank $(n-1, m-1)$, $(n-2, m+1)$ and $(n+1, m-2)$ by constructing respectively the following products and summing over repeated indices:

$$T_{j_1 \cdots j_m}^{i_1 \cdots i_n}\; \delta_{i_a}^{\,j_b} \quad , \tag{12.11}$$

---

* Tensors whose components are unaltered under the transformations are called isotropic.

$$T^{i_1 \cdots i_n}_{j_1 \cdots j_m} \epsilon_{j_{m+1} i_a i_{a'}} \tag{12.12}$$

and

$$T^{i_1 \cdots i_n}_{j_1 \cdots j_m} \epsilon^{i_{n+1} j_b j_{b'}} , \tag{12.13}$$

where a and a' are two different integers between 1 and n, and likewise b and b' are different integers between 1 and m, assuming that n and m are both $\geqslant 1$ in (12.11), $n \geqslant 2$ in (12.12) and $m \geqslant 2$ in (12.13).

Definition: A tensor T of rank (n, m) is called reducible if through contraction a new nonzero tensor T' of rank (n', m') can be formed with

$$n' + m' < n + m \ ;$$

otherwise T is irreducible.

Because of (12.12)-(12.13), an irreducible tensor $T^{i_1 \cdots i_n}_{j_1 \cdots j_m}$ must be symmetric with respect to any pair $(i_a, i_{a'})$ or $(j_b, j_{b'})$. Furthermore, in view of (12.11), T must also satisfy the trace condition

$$T^{i_1 \cdots i_n}_{j_1 \cdots j_m} \delta^{j_1}_{i_1} = 0 \ . \tag{12.14}$$

Because of the symmetry and trace conditions, the components of an irreducible tensor may not all be independent.

2. Representations

Consider a tensor of rank (n, m). Let the number of its linearly-independent components be d, known as the dimension. From such a tensor, we can select d linearly-independent components $\phi_1, \phi_2, \cdots, \phi_d$ and write

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_d \end{pmatrix} \ .$$

Under (12.2), $q \to v\,q$, the transformation (12.7) can be written in terms of ϕ as

$$\phi \to V \cdot \phi \tag{12.15}$$

where V is a $d \times d$ matrix. Because the V's clearly satisfy the same algebra as the v's, we regard $\{V\}$ as forming a representation, denoted by $\textcircled{d}$, of the SU_3 group $\{v\}$. Furthermore, the representation is called <u>irreducible</u> if the tensor is; otherwise it is reducible.

The following is a list of the low-ranking irreducible tensors:

Irreducible tensor	Rank	Representation
1	$(0, 0)$	$\textcircled{1}$
T^i	$(1, 0)$	$\textcircled{3}$
T_a	$(0, 1)$	$\textcircled{$\bar{3}$}$
T^i_j	$(1, 1)$	$\textcircled{8}$
T^{ij}	$(2, 0)$	$\textcircled{6}$
T_{ab}	$(0, 2)$	$\textcircled{$\bar{6}$}$
T^{ijk}	$(3, 0)$	$\textcircled{10}$
T_{abc}	$(0, 3)$	$\textcircled{$\bar{10}$}$
T^{ij}_{ab}	$(2, 2)$	$\textcircled{27}$

TABLE 12.1

where the numbers in the last column denote the dimensions. We note that both T^i and T_a are of dimension three. However, because

of the difference in their transformation properties, these form two different representations, called (3) and ($\bar{3}$) . [For further analysis, see the discussions given on page 261.] Due to the symmetry condition, $T^{ij} = T^{ji}$ and $T_{ab} = T_{ba}$, both representations are of dimension six, but each is different and hence they are labelled (6) and ($\bar{6}$) . Because of the constraint $T^{i}_{i} = 0$, the irreducible tensor T^{i}_{a} is of dimension $3^2 - 1 = 8$. For a symmetric tensor T^{ijk}, there are three independent components of the type T^{111}, six of the type T^{112} and one of the type T^{123}, making a total of ten and forming the representation (10) . Similarly, the irreducible tensor T_{abc} is also of dimension ten, but labelled ($\overline{10}$) because it forms a representation different from that given by T^{ijk} . In like manner, it is straightforward to find that the irreducible tensor T_{ab}^{ij} is of dimension (27) .

3. Decomposition of (8) x (8)

Let A^{i}_{j} and B^{i}_{j} be both of the (8) representation. Hence,

$$A^{i}_{i} = B^{i}_{i} = 0 \quad . \tag{12.16}$$

Consider the product (8) x (8) :

$$A^{i}_{a} B^{j}_{b} \quad , \tag{12.17}$$

which is a reducible tensor of rank (2, 2). Our task is to form linear functions of its $8^2 = 64$ components so that they become irreducible.

(i) The sum

$$S = A^{i}_{a} B^{a}_{i} \tag{12.18}$$

is clearly an irreducible tensor of rank (0, 0) .

(ii) The tensor

$$F_a^{\ i} = A_j^{\ i} B_a^{\ j} - B_j^{\ i} A_a^{\ j} \tag{12.19}$$

is traceless and therefore forms an irreducible 8-dimensional representation. Likewise

$$D_a^{\ i} = A_j^{\ i} B_a^{\ j} + B_j^{\ i} A_a^{\ j} - \tfrac{2}{3} \delta_j^{\ i} S \tag{12.20}$$

forms another irreducible 8-dimensional representation. Since $F_a^{\ i}$ is antisymmetric in A and B, while $D_a^{\ i}$ is symmetric, these are two different functions, although both are (8) .

(iii) The symmetric tensor

$$T^{ijk} = A_a^{\ i} B_b^{\ j} \epsilon^{abk} + \text{all terms formed by permuting } ijk$$

gives an irreducible representation (10) . Similarly, by interchanging superscripts and subscripts, we can form a representation $(\overline{10})$ in terms of the symmetric tensor

$$\overline{T}_{abc} = A_a^{\ i} B_b^{\ j} \epsilon_{ijc} + \text{all terms formed by permuting } abc .$$

(iv) By symmetrizing (12.17) with respect to (i, j) and (a, b), we can first form

$$R_{ab}^{\ ij} = A_a^{\ i} B_b^{\ j} + A_a^{\ j} B_b^{\ i} + A_b^{\ i} B_a^{\ j} + A_b^{\ j} B_a^{\ i} ,$$

which satisfies, on account of (12.18) and (12.20),

$$R_{ab}^{\ ib} = D_a^{\ i} + \tfrac{2}{3} \delta_a^{\ i} S .$$

We then construct the irreducible tensor

$$I_{ab}^{\ ij} = R_{ab}^{\ ij} - \tfrac{1}{5} (\delta_a^{\ i} D_b^{\ j} + \delta_a^{\ j} D_b^{\ i} + \delta_b^{\ i} D_a^{\ j} + \delta_b^{\ j} D_a^{\ i}) - \tfrac{1}{6} (\delta_a^{\ i} \delta_b^{\ j} + \delta_a^{\ j} \delta_b^{\ i}) S ,$$

which forms an irreducible representation (27) .

Putting together (i)-(iv), we can write

$$(8) \times (8) = (1) + (8) + (8) + (10) + (\overline{10}) + (27) \quad .$$

It is not difficult to extend the above considerations to representations of higher dimensions.

4. Some further properties

Consider an infinitesimal SU_3 transformation which contains, because of $\det v = 1$, $3^2 - 1 = 8$ independent real infinitesimal quantities $\epsilon_1, \epsilon_2, \cdots, \epsilon_8$:

$$v = 1 + \tfrac{1}{2} i \epsilon_\ell \lambda_\ell \tag{12.21}$$

where the λ_ℓ's are eight 3×3 matrices and 1 denotes the unit matrix. Since

$$v v^\dagger = 1 + \tfrac{1}{2} i \epsilon_\ell (\lambda_\ell - \lambda_\ell^\dagger) + O(\epsilon^2) = 1 \quad ,$$

we have

$$\lambda_\ell = \lambda_\ell^\dagger \quad .$$

These eight Hermitian matrices $\lambda_1, \lambda_2, \cdots, \lambda_8$ are called the generators of the SU_3 group; they play the same role as the three Pauli matrices for the SU_2 group. It is customary to write

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \tag{12.22}$$

$$\lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad .$$

The trace, commutator and anticommutator of two λ matrices are

$$\begin{aligned} \mathrm{tr}\,(\lambda_\ell\,\lambda_m) &= 2\delta_{\ell m}\ , \\ [\lambda_\ell\ ,\ \lambda_m\,] &= 2i\,f_{\ell mn}\,\lambda_n \\ \text{and} \\ \{\lambda_\ell\ ,\ \lambda_m\} &= \frac{4}{3}\,\delta_{\ell m} + 2d_{\ell mn}\,\lambda_n \end{aligned} \tag{12.23}$$

where $f_{\ell mn}$ is completely antisymmetric in its indices while $d_{\ell mn}$ is completely symmetric. The nonzero elements of $f_{\ell mn}$ and $d_{\ell mn}$ are as follows:

ℓmn	$f_{\ell mn}$	ℓmn	$d_{\ell mn}$
123	1	118	$\sqrt{\frac{1}{3}}$
147	$\frac{1}{2}$	146	$\frac{1}{2}$
156	$-\frac{1}{2}$	157	$\frac{1}{2}$
246	$\frac{1}{2}$	228	$\sqrt{\frac{1}{3}}$
257	$\frac{1}{2}$	247	$-\frac{1}{2}$
345	$\frac{1}{2}$	256	$\frac{1}{2}$
367	$-\frac{1}{2}$	338	$\sqrt{\frac{1}{3}}$
458	$\frac{1}{2}\sqrt{3}$	344	$\frac{1}{2}$
678	$\frac{1}{2}\sqrt{3}$	355	$\frac{1}{2}$
		366	$-\frac{1}{2}$
		377	$-\frac{1}{2}$
		448	$-1/(2\sqrt{3})$
		558	$-1/(2\sqrt{3})$
		668	$-1/(2\sqrt{3})$
		778	$-1/(2\sqrt{3})$
		888	$-\sqrt{\frac{1}{3}}$.

(12.24)

Let $\{v\}$ and $\{\bar{v}\}$ be two 3×3 irreducible representations

of the SU_3 group. If for every $\bar{v}$ there is a v such that

$$\bar{v} = v_o \, v \, v_o^\dagger \tag{12.25}$$

where v_o is a fixed element of $\{v\}$, then all the $\bar{v}$'s also belong to $\{v\}$; therefore these two representations are regarded as the same, otherwise not. In Table 12.1 the representation associated with the irreducible tensor T^i is $\{v\}$, while that associated with T_a is $\{v^*\}$. We denote these two representations as $\textcircled{3}$ and $\textcircled{\bar{3}}$ because, as we shall now prove, they are not the same.

Proof. If we were to assume that they are, (12.25) would become

$$v^* = v_o \, v \, v_o^\dagger \quad . \tag{12.26}$$

Consider now an infinitesimal v of the form (12.21), whose complex conjugation is

$$v^* = 1 - \tfrac{1}{2} i \, \lambda_\ell^* \, \epsilon_\ell \quad .$$

Upon substituting this expression, together with (12.21), into (12.26) we find

$$\lambda_\ell^* = - v_o \, \lambda_\ell \, v_o^\dagger \quad . \tag{12.27}$$

From (12.22) we see that

$$\lambda_\ell^* = \pm \lambda_\ell \tag{12.28}$$

where the $\pm$ sign $\equiv \eta_\ell$ is given by

$$\eta_\ell = \begin{cases} + & \text{for } \ell = 1, 3, 4, 6, 8 \\ - & \text{for } \ell = 2, 5, 7 \end{cases} \tag{12.29}$$

and therefore (12.27) becomes

$$- \eta_\ell \, \lambda_\ell = v_o \, \lambda_\ell \, v_o^\dagger \tag{12.30}$$

in which the repeated index is not summed over, but can be 1 or 2, $\cdots$,

or 8 . By multiplying the last equation in (12.23) on the left by v_o and on the right by $v_o^\dagger$, we see that it is invariant under the change $\lambda_\ell \to v_o \lambda_\ell v_o^\dagger$; thus on account of (12.30) we can derive

$$- \eta_\ell \eta_m \eta_n d_{\ell mn} = d_{\ell mn} \tag{12.31}$$

where as in (12.30) the repeated indices refer to fixed numbers, not summed over. By using (12.24) and (12.29), we find that this is wrong for every nonzero $d_{\ell mn}$. Therefore, the representations $\{v\}$ and $\{v^*\}$ are different.

Exercise. Establish first the Jacobi identity

$$[\lambda_\ell, [\lambda_m, \lambda_n]] + [\lambda_m, [\lambda_n, \lambda_\ell]] + [\lambda_n, [\lambda_\ell, \lambda_m]] = 0$$

and then derive

$$f_{\ell\ell' k} f_{mn\ell'} + f_{m\ell' k} f_{n\ell\ell'} + f_{n\ell' k} f_{\ell m\ell'} = 0 \ .$$

5. Excursion to other groups

(i) SU_2 group

In the case of the SU_2 group $\{s\}$, on account of (11.42) we have for every s

$$s^* = s_o s s_o^\dagger \tag{12.32}$$

where s_o is a fixed element given by

$$s_o = e^{i\frac{1}{2}\pi\tau_2} = i\tau_2 \ .$$

Consequently, representations $\{s\}$ and $\{s^*\}$ are the same (sometimes referred to as equivalent); hence, there exists only one two-dimensional representation of the SU_2 group.

The notion of tensors and representation can be applied equally well to the SU_2 group. We call $T_{a_1 \cdots a_n}$ the component of a tensor

of rank n if its transformation law is

$$T_{a_1 a_2 \cdots a_n} \to s_{a_1 b_1} s_{a_2 b_2} \cdots s_{a_n b_n} T_{b_1 b_2 \cdots b_n} \tag{12.33}$$

where a_i and b_j can be 1 or 2. Because of (12.32), the rank is now characterized by only one number. Since $\det s = 1$, one sees that

$$\epsilon_{ab} \equiv \begin{cases} 1 & \text{for } a = 1,\ b = 2, \\ -1 & \text{for } a = 2,\ b = 1, \\ 0 & \text{otherwise} \end{cases} \tag{12.34}$$

is an isotropic tensor of rank 2.

Similarly to (12.12)-(12.13), if a tensor T is not completely symmetric in its indices, say $T_{a_1 a_2 a_3 \cdots a_n} \neq T_{a_2 a_1 a_3 \cdots a_n}$, then by constructing

$$\epsilon_{a_1 a_2} T_{a_1 a_2 \cdots a_n}$$

we can form a nonzero tensor of lower rank $n - 2$. Such a tensor T is called reducible; otherwise it is irreducible. Again each tensor gives a representation of the group, and the representation is irreducible if the tensor is. A tensor of rank n has 2^n components which may not all be independent. The dimension d is defined to be the number of its linearly-independent components.

Exercise. Prove the following list of irreducible tensors and representations of the SU_2 group:

	Irreducible tensor	Rank	Dimension of representation
	1	0	1
	T_a	1	2
symmetric	T_{ab}	2	3
completely symmetric	T_{abc}	3	4
	. . . . .		
completely symmetric	$T_{a_1 a_2 \cdots a_n}$	n	n + 1 .

This classification of irreducible representations is identical to the usual one in terms of angular-momentum states, provided $n = 2j$ where j is the total angular-momentum quantum number.

(ii) SO_3 group

If we limit ourselves to the angular-momentum j=integer states, the corresponding group becomes SO_3, which comprises all 3×3 real, orthogonal matrices u; i.e.,

$$u = u^*, \quad u\tilde{u} = 1 \quad \text{and} \quad \det u = 1 \ . \tag{12.35}$$

Each u can be viewed as a three-dimensional rotation of the position vector $\vec{r} = (x_1, x_2, x_3)$:

$$x_i \rightarrow u_{ij} x_j$$

where the subscripts i and j vary from 1 to 3. The component of a tensor T of rank n now transforms according to

$$T_{i_1 i_2 \cdots i_n} \rightarrow u_{i_1 j_1} u_{i_2 j_2} \cdots u_{i_n j_n} T_{j_1 j_2 \cdots j_n} \ .$$

Because $u\tilde{u} = 1$ and $\det u = 1$, one sees that δ_{ij} and ϵ_{ijk} respectively form tensors of rank 2 and 3, where δ_{ij} is a Kronecker symbol and ϵ_{ijk} is defined by (3.4). A tensor is called reducible if it is either not totally symmetric with respect to its indices, say $T_{i_1 i_2 i_3 \cdots i_n} \neq T_{i_2 i_1 i_3 \cdots i_n}$, or it has nonzero trace. In the former we can form a nonzero tensor of rank $n - 1$ by constructing

$$T_{i_1 i_2 i_3 \cdots i_n} \epsilon_{i_1 i_2 j} \ ;$$

while in the latter we can form a tensor of rank $n-2$ through the contraction

$$T_{i_1 i_2 \cdots i_n} \delta_{i_a i_b}$$

where a and b can be any two different integers between 1 and n. The notions, previously discussed, of dimension and representation and its reducibility can be straightforwardly applied to the present case.

Exercise. Establish the following list of irreducible tensors and representations of the SO_3 group:

Irreducible tensor	Rank	Dimension of representation
1	0	1
T_i	1	3
T_{ij}	2	5
T_{ijk}	3	7
. . . . .		
$T_{i_1 \cdots i_\ell}$	ℓ	$2\ell + 1$.

Show that the irreducible representations can be chosen to be precisely the familiar spherical harmonics $Y_{\ell m}$ discussed in (1.38)-(1.40).

12.2 Hadron States and Their Flavor and Color Symmetries

It is convenient to denote the q^i's of (12.1) as the specific up, down and strange quarks written respectively as

$$q^1 = u \ , \quad q^2 = d \quad \text{and} \quad q^3 = s \ . \tag{12.36}$$

The approximate invariance under the transformation $q \to vq$, given by (12.2), will be referred to as the SU_3 flavor symmetry. The dynamic properties of these quarks will be examined later. Here we only use their transformation properties. The q^i's form the irreducible representation (3) ; their antiparticles $q_1 = \bar{u}$, $q_2 = \bar{d}$ and $q_3 = \bar{s}$ form the representation ($\bar{3}$) . The isospin SU_2 is a subgroup of SU_3 with $u = \uparrow$, $d = \downarrow$ forming an isospin doublet and s an isospin singlet.

In 3×3 matrix form, each 2×2 isospin transformation matrix s corresponds to

$$v = \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix} \tag{12.37}$$

where 0 stands for either a 2×1 or a 1×2 null matrix.

1. Pseudoscalar octet

The eight pseudoscalar mesons listed in Table 11.1 on page 239 are regarded as forming an SU_3 octet M_j^i with

$$M_i^i = 0 \ , \tag{12.38}$$

which has the same transformation properties as the quark-antiquark system:

$$M_j^i \sim q^i q_j - \tfrac{1}{3} \delta_j^i q^k q_k \ . \tag{12.39}$$

By comparing the isospin properties of the pseudoscalar mesons with those of the quarks, we find

$$\begin{aligned} &\pi^+ \sim u\bar{d} \ , \quad \pi^- \sim d\bar{u} \ , \\ &\pi^o \sim \frac{1}{\sqrt{2}} (u\bar{u} - d\bar{d}) \ , \\ &K^+ \sim u\bar{s} \ , \quad K^- \sim s\bar{u} \ , \\ &K^o \sim d\bar{s} \ , \quad \bar{K}^o \sim s\bar{d} \ , \\ &\eta^o \sim \frac{1}{\sqrt{6}} (2s\bar{s} - u\bar{u} - d\bar{d}) \ . \end{aligned} \tag{12.40}$$

Here the $\sim$ indicates that its two sides have the same SU_3 transformation properties. Just as in (11.45)-(11.46), the isospin transformation properties can be exhibited by writing

$$u = \uparrow \ , \quad d = \downarrow \tag{12.41}$$

for the quark isodoublet, and

$$\bar{d} = \uparrow \text{ and } -\bar{u} = \downarrow \tag{12.42}$$

for the antiquark states. Thus, the first expression in (12.40) gives

$$\pi^+ \sim \uparrow\bar{\uparrow}$$

where the first $\uparrow$ on the left refers to the quark and the second to the antiquark. Likewise $\pi^- \sim -\downarrow\bar{\downarrow}$ and $\pi^o \sim -\frac{1}{\sqrt{2}}(\uparrow\bar{\downarrow}+\downarrow\bar{\uparrow})$ which are the familiar $I_z = -1$ and 0 states of an isospin triplet. Notice that in π^o the relative sign between $\uparrow\bar{\downarrow}$ and $\downarrow\bar{\uparrow}$ is $+$, while according to (12.40) that between $u\bar{u}$ and $d\bar{d}$ in π^o is $-$, which leads to, on account of (12.39),

$$\pi^o = \frac{1}{\sqrt{2}}(M_1^{\ 1} - M_2^{\ 2}) \quad . \tag{12.43}$$

Because of the trace condition (12.38), from the three diagonal matrix elements $M_1^{\ 1}$, $M_2^{\ 2}$ and $M_3^{\ 3}$ there are only two linearly independent components. One is π^o and the other is η^o. Apart from an arbitrary overall phase factor, by using (12.39) one sees that the expression for η^o is uniquely determined by its orthogonality relation with π^o and its normalization condition. The resulting expression is given by the last formula in (12.40), which may also be written as

$$\eta^o = \frac{1}{\sqrt{6}}(-M_1^{\ 1} - M_2^{\ 2} + 2\,M_3^{\ 3}) \quad . \tag{12.44}$$

To express $M_1^{\ 1}$, $M_2^{\ 2}$ and $M_3^{\ 3}$ in terms of π^o and η^o, we may temporarily refrain from using (12.38), but formally regard

$$M_i^{\ i} = M_1^{\ 1} + M_2^{\ 2} + M_3^{\ 3} \tag{12.45}$$

as an entity. Solving (12.43)-(12.45) for $M_1^{\ 1}$, we find

$$M_1^{\ 1} = \frac{\pi^o}{\sqrt{2}} - \frac{\eta^o}{\sqrt{6}} + \tfrac{1}{3}\,M_i^{\ i}$$

which, on account of (12.38), gives

$$M_1^{\ 1} = \frac{\pi^o}{\sqrt{2}} - \frac{\eta^o}{\sqrt{6}} \quad . \tag{12.46}$$

Likewise, we can solve M_2^2 and M_3^3. Thus, from (12.40) it follows that the pseudoscalar octet matrix is

$$M = (M_j^i) = \begin{pmatrix} \frac{\pi^o}{\sqrt{2}} - \frac{\eta^o}{\sqrt{6}} & \pi^+ & K^+ \\ \pi^- & -\frac{\pi^o}{\sqrt{2}} - \frac{\eta^o}{\sqrt{6}} & K^o \\ K^- & \bar{K}^o & \frac{2\eta^o}{\sqrt{6}} \end{pmatrix}. \tag{12.47}$$

2. Baryon spin-$\frac{1}{2}$ octet and spin-$\frac{3}{2}$ decuplet

The eight spin-$\frac{1}{2}$ baryons listed in Table 11.1 are also regarded as forming an SU_3 octet B_j^i. By matching I, I_z and Y, we replace $\pi \to \Sigma$, $\eta^o \to \Lambda^o$, $K^+ \to p$, $K^o \to n$, $K^- \to \Xi^-$ and $\bar{K}^o \to \Xi^o$; the meson octet becomes the baryon octet and (12.47) becomes

$$B = (B_j^i) = \begin{pmatrix} \frac{\Sigma^o}{\sqrt{2}} - \frac{\Lambda^o}{\sqrt{6}} & \Sigma^+ & p \\ \Sigma^- & -\frac{\Sigma^o}{\sqrt{2}} - \frac{\Lambda^o}{\sqrt{6}} & n \\ \Xi^- & \Xi^o & \frac{2\Lambda^o}{\sqrt{6}} \end{pmatrix}. \tag{12.48}$$

In the Table of Particle Properties, the low-lying spin-$\frac{3}{2}$ baryon resonances are $\Delta(1232)$, $\Sigma^*(1385)$, $\Xi^*(1530)$ and $\Omega^-(1672)$. These will be identified as forming an SU_3 decuplet D^{ijk}. The assignments can be done best in terms of the quark model.

Because of the half-integer nature of the baryons, in a quark model they must be composites of an odd number of quarks and antiquarks. The simplest way is to regard them as bound states of three quarks, each a spin-$\frac{1}{2}$ fermion. If we exhibit only the SU_3-flavor

indices, as in (12.39), then we may write

$$B_j^{\ i} \sim q^i q^a q^b \epsilon_{abj} - \tfrac{1}{3}\delta_j^{\ i} q^k q^a q^b \epsilon_{abk}$$

and

$$D^{ijk} \sim \text{symmetrized } q^i q^j q^k \ . \tag{12.49}$$

The assignments of the decuplet D^{ijk}, in the notation of (12.36), are

$$\begin{aligned}
\Omega^- &= D^{333} \sim sss \ , \\
\Xi^{*0} &\sim \sqrt{\tfrac{1}{3}}\,(uss + sus + ssu) \ , \\
\Xi^{*-} &\sim \sqrt{\tfrac{1}{3}}\,(dss + sds + ssd) \ , \\
\Sigma^{*+} &\sim \sqrt{\tfrac{1}{3}}\,(suu + usu + suu) \ , \\
\Sigma^{*0} &\sim (sud + sdu + usd + dsu + uds + dus)/\sqrt{6} \ , \\
\Sigma^{*-} &\sim \sqrt{\tfrac{1}{3}}\,(sdd + dsd + sdd) \ , \\
\Delta^{++} &\sim uuu \ , \\
\Delta^{+} &\sim \sqrt{\tfrac{1}{3}}\,(duu + udu + uud) \ , \\
\Delta^{0} &\sim \sqrt{\tfrac{1}{3}}\,(udd + dud + ddu) \ , \\
\Delta^{-} &\sim ddd \ ,
\end{aligned} \tag{12.50}$$

in which the different orders of quarks are regarded as having some different kinematic attributes. For the moment, if one wishes, one may simply choose these different attributes to be, say, different momenta $\vec{p}$, $\vec{p}'$, $\vec{p}''$ with $\vec{p} + \vec{p}' + \vec{p}'' = 0$. Hence, e.g., in Ξ^{*0} one may view the first term uss as referring to u having $\vec{p}$, the middle s having $\vec{p}'$ and the final s having $\vec{p}''$; the two other terms sus and ssu are generated by permuting the SU_3 indices, but keeping $\vec{p}$, $\vec{p}'$, $\vec{p}''$ fixed. This, then, explains the $\sqrt{\tfrac{1}{3}}$ factor in Ξ^{*0} so that it is normalized in the same way as Ω^-. [For further

details, see (12.51) and (12.54)-(12.56) below.] The actual dynamics of these bound states will be analyzed later in Chapter 20. Here we are interested in their purely kinematic aspects. From (12.50) we see that the charges of u, d and s must be $\frac{2}{3}$, $-\frac{1}{3}$ and $-\frac{1}{3}$ respectively. Their spins are $\frac{1}{2}$, and hence they are fermions.

As we shall show later, it is a reasonable approximation to assume further that each quark is in the same lowest-energy s-orbit of a common potential with no mutual interaction. This leads to an immediate problem: Consider, for example, the spin-$\frac{3}{2}$ baryon $\Omega^- \sim sss$. After we sum over the different quark momenta, because of the s-orbit assumption, the total orbital wave function is completely symmetric. So is the total spin wave function because all three $\frac{1}{2}$-quark spins are lined up to form a total $\frac{3}{2}$-spin. The Fermi statistics of quarks would make it impossible for them to be the component of a completely symmetric tensor D^{ijk} unless we assume that the quarks have some other degrees of freedom. For this reason, we shall assume that the quarks do have another degree of freedom*, called <u>color</u>. Each quark, u or d or s $\cdots$, has three varieties u(c) or d(c) or s(c) $\cdots$, with $c = 1, 2, 3$. [Sometimes, for more colorful designations one might choose, e.g., c = red, white and blue.] For example, in the second expression of (12.49), the baryon state with total spin = $\frac{3}{2}$ and its z-component also = $\frac{3}{2}$, assumes then the form

$$D^{ijk} \sim q_{\uparrow}^{i}(c)\, q_{\uparrow}^{j}(c')\, q_{\uparrow}^{k}(c'')\, \epsilon_{cc'c''} \tag{12.51}$$

where c, c', c'' are the color indices, $\epsilon_{cc'c''}$ is the totally antisymmetric tensor given by (3.4), the subscripts $\uparrow$ and $\downarrow$ denote

---

* O. W. Greenberg, Phys. Rev. Lett. <u>13</u>, 598 (1964).

the spin-up and -down configurations of the quark, and $q_{\uparrow}^{i}(c)$ represents the corresponding anticommuting quark-field operator. Consequently, the interchange of i and j on the righthand side of (12.51) results in two minus signs: one from the anticommutation of the fermion operators and the other from $\epsilon_{cc'c''}$. Hence, the symmetry of D^{ijk} is insured.

Now, let us examine other hadron states with this new degree of freedom included. Consider, e.g., (12.40); instead of $\pi^{+} \sim u\bar{d}$, we might write $u(c)\,\bar{d}(c')$ which could lead to $3^2 = 9$ possible meson states because of different values of c and c'. In order to avoid this difficulty, we postulate that all interactions satisfy exact SU_3-color symmetry and all observed hadron states are color singlets. Hence, π^{+} is represented by

$$\pi^{+} \sim u(c)\,\bar{d}(c) \tag{12.52}$$

with the repeated color index c summed over from 1 to 3 so that the state becomes a color singlet, as in (12.51). The same rule applies to all other meson and baryon states; hence (12.39) and (12.49) become

$$\begin{aligned} M_j^{\,i} &\sim q^{i}(c)\,q_j(c) - \tfrac{1}{3}\,\delta_j^{\,i}\,q^{k}(c)\,q_k(c)\ , \\ B_j^{\,i} &\sim [\,q^{i}(c)\,q^{a}(c')\,q^{b}(c'')\,\epsilon_{abj} \\ &\qquad - \tfrac{1}{3}\,\delta_j^{\,i}\,q^{k}(c)\,q^{a}(c')\,q^{b}(c'')\,\epsilon_{abk}\,]\,\epsilon_{cc'c''} \end{aligned} \tag{12.53}$$

and

$$D^{ijk} \sim q^{i}(c)\,q^{j}(c')\,q^{k}(c'')\,\epsilon_{cc'c''}$$

where, for clarity, the spin-dependence is not exhibited.

As noted before, the indices that differentiate the u, d and s quarks are referred to as flavor indices. For example, D^{ijk} has ten physically different flavor configurations corresponding to different

choices of superscript:

$$\begin{aligned}
\Omega^- &= D^{333} , \\
\Xi^{*0} &= \sqrt{\tfrac{1}{3}}\,(D^{133} + D^{313} + D^{331}) , \\
\Xi^{*-} &= \sqrt{\tfrac{1}{3}}\,(D^{233} + D^{323} + D^{332}) , \\
\Sigma^{*+} &= \sqrt{\tfrac{1}{3}}\,(D^{311} + D^{131} + D^{113}) , \\
\Sigma^{*0} &= (D^{312} + D^{321} + D^{132} + D^{231} + D^{123} + D^{213})/\sqrt{6} , \\
\Sigma^{*-} &= \sqrt{\tfrac{1}{3}}\,(D^{322} + D^{232} + D^{223}) , \\
\Delta^{++} &= D^{111} , \\
\Delta^{+} &= \sqrt{\tfrac{1}{3}}\,(D^{211} + D^{121} + D^{112}) , \\
\Delta^{0} &= \sqrt{\tfrac{1}{3}}\,(D^{122} + D^{212} + D^{221}) , \\
\Delta^{-} &= D^{222} ,
\end{aligned} \tag{12.54}$$

where, for the z-component spin $= \frac{3}{2}$ states,

$$\begin{aligned}
D^{333} &\sim s_{\uparrow}(c)\, s_{\uparrow}(c')\, s_{\uparrow}(c'')\, \epsilon_{cc'c''} , \\
D^{133} &\sim u_{\uparrow}(c)\, s_{\uparrow}(c')\, s_{\uparrow}(c'')\, \epsilon_{cc'c''} , \\
D^{313} &\sim s_{\uparrow}(c)\, u_{\uparrow}(c')\, s_{\uparrow}(c'')\, \epsilon_{cc'c''} ,
\end{aligned} \tag{12.55}$$

etc., in accordance with the notations of (12.36) and (12.51). Through space rotations, decuplet baryons with different spin components can be readily generated. Because D^{ijk} is symmetric with respect to i, j and k, we have from (12.54)

$$\begin{aligned}
D^{133} = D^{313} = D^{331} &= \sqrt{\tfrac{1}{3}}\ \Xi^{*0} , \\
D^{311} = D^{131} = D^{113} &= \sqrt{\tfrac{1}{3}}\ \Sigma^{*+} , \\
D^{211} = D^{121} = D^{112} &= \sqrt{\tfrac{1}{3}}\ \Delta^{+}
\end{aligned} \tag{12.56}$$

and similar expressions for Ξ^{*-}, Σ^{*0}, Σ^{*-} and Δ^0.

Remarks. For the lowest-energy baryon states, the three quarks are all in the same s-state. Hence, the total angular momentum equals the total spin, which can be $\frac{1}{2}$ or $\frac{3}{2}$; the former gives the spin-$\frac{1}{2}$ octet and the latter the spin-$\frac{3}{2}$ decuplet. For a similarly constructed meson system of the s-state quark-antiquark pair, the total spin can be 0 or 1; the former corresponds to the pseudoscalar octet plus $\eta'(958)$, and the latter to the vector nonet: $\rho(770)$, $\omega(783)$, $K^*(892)$ and $\phi(1020)$. Further discussions will be given in Chapter 20.

12.3 Mass Formulas

Since π, K and η differ in their masses by a few hundred MeV, as do some of the different members of the baryon octet or decuplet, the strong interaction can at best be only approximately SU_3-flavor symmetric. In contrast, as mentioned before, the SU_3-color symmetry is assumed to be exact. In this section we shall deal only with SU_3-flavor transformations; hence the word "flavor" will be omitted for notational convenience.

Let us decompose

$$H_{st} = H_{sym} + H_{asym} \tag{12.57}$$

where H_{sym} is SU_3-invariant, while H_{asym} is not.

1. H_{asym} and the spurion formulation

We may envisage the expression of H_{asym} in terms of the quark field q^i given by (12.1). In order to conserve baryon number, the simplest form would be a linear function of the quadratic expression $q^{i\dagger} q^j$; i.e., H_{asym} transforms like an irreducible tensor of rank (1, 1). Conservation of charge and isospin then requires H_{asym} to

transform as

$$H_{asym} \sim S_3^3 \tag{12.58}$$

which is the $i = j = 3$ member of an octet S_j^i. The precise form of S_j^i is immaterial, since in this section we are interested only in the SU_3 transformation properties of H_{asym}. [As a concrete example, we may assume S_j^i to be given by

$$S_j^i = q_j \beta q^i - \tfrac{1}{3} \delta_j^i q_k \beta q^k$$

where β is the Dirac matrix given by (3.10), and q_j is the Hermitian conjugate of the quark field q^j.] The energy E of a hadron h is then given by the diagonal matrix element

$$E_h = \langle h \mid H_{st} \mid h \rangle \ . \tag{12.59}$$

In the following, we shall assume (12.57); furthermore, H_{asym} is supposed to be much weaker than H_{sym}, and therefore it can be regarded as a perturbation. Neglecting second-order effects of H_{asym} in (12.59), we need the state vector $\mid h \rangle$ only to the accuracy of the zeroth order; i.e., $\mid h \rangle$ satisfies

$$H_{sym} \mid h \rangle = \gamma \mid h \rangle \tag{12.60}$$

where γ is the eigenvalue. Hence, (12.59) becomes

$$E_h = \gamma + \langle h \mid H_{asym} \mid h \rangle + O(H_{asym}^2) \ . \tag{12.61}$$

To incorporate the SU_3 transformation property (12.58) of H_{asym} into the energy calculation, we introduce the "spurion" formulation of G. Wentzel.* Let us examine a typical matrix element $\langle h' \mid H_{asym} \mid h \rangle$. Both $\mid h \rangle$ and $\mid h' \rangle$ are eigenvectors of H_{sym}

* G. Wentzel, in High Energy Nuclear Physics (Proceedings of the Rochester Conference), ed. J. Ballam et al. (New York, Interscience Publishers, 1956), VIII 15.

and therefore belong to some irreducible representations of SU_3. We then consider a hypothetical SU_3-conserving transition

$$h \rightarrow h' + S_j^i \tag{12.62}$$

where S_j^i denotes the spurion, which transforms as an irreducible SU_3 octet representation, and carries zero 4-momentum, zero charge and even parity; in addition, we are interested only in the final state $i = j = 3$ of the spurion. It is clear that so far as the SU_3 transformations are concerned, the matrix element $< h' \mid H_{asym} \mid h >$, with H_{asym} given by (12.58), has the same properties as that of the hypothetical SU_3-conserving transition amplitude for (12.61). Emitting a spurion with SU_3 symmetry conserved is identical to having an appropriate SU_3-violating amplitude, but without the spurion. In the following we shall see how to derive various mass formulas* by using the spurion.

2. Octet mass formulas

We first discuss the matrix element (12.59) for the baryon octet $h = B_j^i$, given by (12.48). In terms of the spurion formulation, the relevant transition (12.62) becomes

$$B_j^i \rightarrow B_b^a + S_3^3 \ .$$

Phenomenologically we may regard H_{asym} as the transition Hamiltonian given by an SU_3-conserving sum of the products of the three octets: the baryon field $B = (B_b^a)$, its Hermitian conjugate $\bar{B} = (\bar{B}_a^b)$ and the spurion field $S = (S_j^i)$. By using (12.19) and (12.20), we see that among such products there can be only two invariants:

---

*M. Gell-Mann, Phys. Rev. 125, 1067 (1962). S. Okubo, Progr. Theor. Phys. 27, 949 (1962).

$$X_i^{\ j}\, S_j^{\ i} \qquad \text{and} \qquad Y_i^{\ j}\, S_j^{\ i} \tag{12.63}$$

where

$$X_i^{\ j} = \bar{B}_a^{\ j}\, B_i^{\ a}$$

and (12.64)

$$Y_i^{\ j} = \bar{B}_i^{\ a}\, B_a^{\ j} \ .$$

The H_{asym} can be written as

$$H_{asym} = (a\, X_i^{\ j} + \beta\, Y_i^{\ j})\, S_j^{\ i} \tag{12.65}$$

where a and β are constants and, as we shall see, the spurion amplitude can be taken to be

$$S_j^{\ i} = \begin{cases} 1 & \text{if } i = j = 3 \ , \\ 0 & \text{otherwise.} \end{cases} \tag{12.66}$$

For simplicity, we do not separate out the trace of $X_j^{\ i}$, $Y_j^{\ i}$ and $S_j^{\ i}$, since this would only result in a redefinition of H_{asym} and H_{sym} in the decomposition (12.57). We note that in the above expression $S_j^{\ i} = 0$ except for $i = j = 3$, which expresses the hypothesis (12.58) in the spurion language; the value $S_3^{\ 3} = 1$ can be chosen without any loss of generality because of the constants a and β in (12.65). Neglecting $O(H_{asym}^{\ 2})$ and combining (12.65)-(12.66) with (12.57) and (12.61), we can write for the baryon octet

$$H_{st} = a\, X_3^{\ 3} + \beta\, Y_3^{\ 3} + \gamma \tag{12.67}$$

where γ, which is the same constant for different members of the octet, is given by (12.60). By taking the Hermitian conjugate of the baryon octet field (12.48), we have (using a bar instead of † for notational clarity)

$$\bar{B} = (\bar{B}_i^{\ j}) = \begin{pmatrix} \dfrac{\bar{\Sigma}^0}{\sqrt{2}} - \dfrac{\bar{\Lambda}^0}{\sqrt{6}} & \bar{\Sigma}^- & \bar{\Xi}^- \\ \bar{\Sigma}^+ & -\dfrac{\bar{\Sigma}^0}{\sqrt{2}} - \dfrac{\bar{\Lambda}^0}{\sqrt{6}} & \bar{\Xi}^0 \\ \bar{p} & \bar{n} & \dfrac{2\bar{\Lambda}^0}{\sqrt{6}} \end{pmatrix} . \tag{12.68}$$

Therefore, on account of (12.64),

$$X_3^3 = \bar{p}p + \bar{n}n + \tfrac{2}{3}\bar{\Lambda}^0\Lambda^0$$
$$\text{and} \qquad Y_3^3 = \bar{\Xi}^-\Xi^- + \bar{\Xi}^0\Xi^0 + \tfrac{2}{3}\bar{\Lambda}^0\Lambda^0 \quad . \tag{12.69}$$

The diagonal matrix elements (12.59) for different members of the baryon octet can now be derived by using (12.67) and (12.69):

$$E_\Lambda = \tfrac{2}{3}(\alpha+\beta)+\gamma \ , \qquad E_\Sigma = \gamma \ ,$$
$$E_\Xi = \beta + \gamma \qquad \text{and} \qquad E_N = \alpha + \gamma$$

where N stands for p or n. We can eliminate α, β and γ among these four equations. The result is

$$3E_\Lambda - 2(E_N + E_\Xi) + E_\Sigma = 0 \quad . \tag{12.70}$$

Identical considerations can be applied to the pseudoscalar octet. Through the replacements $\Lambda \to \eta$, $N \to K$, $\Xi \to \bar{K}$ and $\Sigma \to \pi$, we derive from (12.70)

$$3E_\eta - 4E_K + E_\pi = 0 \quad . \tag{12.71}$$

Both formulas (12.70)-(12.71) are valid only to the first order in H_{asym}.

In the evaluation of $\langle h \mid H_{st} \mid h \rangle$, we may assume the hadron state to be one with momentum $\vec{k}$. Hence the energy E_h is related to $\vec{k}$ and the hadron mass m_h by

$$E_h = \sqrt{\vec{k}^2 + m_h^2} \quad ; \tag{12.72}$$

its variation is

$$\delta E_h = \frac{\delta(m_h^2)}{2E_h} \quad . \tag{12.73}$$

By substituting (12.73) into (12.70) and (12.71), we obtain the mass formulas, which agree quite well with the experimental data:

$$3m_\Lambda^2 - 2(m_N^2 + m_\Xi^2) + m_\Sigma^2 = 0 \qquad (12.74)$$

and

$$3m_\eta^2 - 4m_K^2 + m_\pi^2 = 0 \ ; \qquad (12.75)$$

in this derivation we employ a reference system in which the hadron energy E_h is much bigger than the mass difference between different octet members, and therefore (12.73) holds. The resulting formulas (12.74)-(12.75) are, of course, independent of the particular reference frame. For the baryon octet we may choose the rest frame, $\vec{k} = 0$, because δm_h is $\ll m_h$; hence (12.73) reduces to $\delta E_h = \delta m_h$ and (12.74) takes on the linear form

$$3m_\Lambda - 2(m_N + m_\Xi) + m_\Sigma = 0 \ . \qquad (12.76)$$

For the meson octet, since m_π is quite a bit smaller than the mass difference $m_K - m_\pi$, it is not possible to linearize (12.75).

Remarks. The two invariants in (12.63), which are formed of the three octets $\bar{B}$, B and S, can be understood by considering first the $(8) \times (8)$ multiplication of the two octets $\bar{B}$ and B. Because

$$(8) \times (8) = (1) + (8) + (8) + (10) + (\overline{10}) + (27) \ ,$$

the product consists of two octets whose components can be derived by using (12.64), and each of which can in turn be combined with S_j^i to form an invariant. This accounts for the two independent constants α and β in (12.65).

3. Decuplet mass formula

From the product of the symmetric decuplet field D^{ijk} and its Hermitian conjugate $\bar{D}_{ijk} = (D^{ijk})^\dagger$, we can form only a single octet

$$Z_j^i = \bar{D}_{jab} D^{iab} - \tfrac{1}{3} \delta_j^i \bar{D}_{abc} D^{abc} \ . \qquad (12.77)$$

By following the same reasoning that led to (12.67), we can write H_{st}

for the baryon decuplet as

$$H_{st} = a + \beta Z_3^3 \tag{12.78}$$

where a and β are constants. Because of (12.54)-(12.56), the above expression becomes

$$H_{st} = a + \beta \left[\tfrac{2}{3} \bar{\Omega}\Omega + \tfrac{1}{3} (\bar{\Xi}^{*0} \Xi^{*0} + \bar{\Xi}^{*-} \Xi^{*-}) - \tfrac{1}{3} (\bar{\Delta}^{++} \Delta^{++} + \bar{\Delta}^{+} \Delta^{+} + \bar{\Delta}^{0} \Delta^{0} + \bar{\Delta}^{-} \Delta^{-}) \right] , \tag{12.79}$$

which leads to $E_\Omega = a + \frac{2}{3}\beta$, $E_{\Xi^*} = a + \frac{1}{3}\beta$, $E_{\Sigma^*} = a$ and $E_\Delta = a - \frac{1}{3}\beta$. Just as in the passage from (12.70) to (12.76), we can arrive at the mass formula

$$m_\Omega - m_{\Xi^*} = m_{\Xi^*} - m_{\Sigma^*} = m_{\Sigma^*} - m_\Delta . \tag{12.80}$$

The reader can easily verify that this mass formula is in good agreement with the experimental values of $\Delta(1232)$, $\Sigma^*(1385)$, $\Xi^*(1530)$ and $\Omega(1672)$ given in the Table of Particle Properties.

Remarks. Considering the fact that m_π is much smaller than the mass differences $m_K - m_\pi$ and $m_\eta - m_\pi$, it seems quite strange that H_{asym} can be treated as a perturbation. A plausible explanation will be given later in Chapter 20.

Problem 12.1. Let $|B_i^a\rangle$ denote the physical spin-$\frac{1}{2}$ baryon octet state and $\langle B_a^i|$ be its Hermitian conjugate.

(i) Prove that if $S_b^j(x)$ is a scalar (i.e., spin-0, parity +) local SU_3-octet operator, then its matrix element between $|B_c^k\rangle$ of 4-momentum k and $|B_i'^a\rangle$ of 4-momentum $k+q$ in the limit $q \to 0$

is given by

$$< B_a'^{\,i} \mid S_b^{\,j}(x) \mid B_c^{\,k} > \;=\; (D_s\, d^{ijk}_{abc} + F_s\, f^{ijk}_{abc})\; u'^{\dagger} \gamma_4 u \qquad (12.81)$$

where u and u' are spinor solutions of (3.26) that have the same spin-momentum configurations as the initial and final physical baryon states, γ_4 is the Dirac matrix given by (3.11), D_s and F_s are constants, and the tensors

$$d^{ijk}_{abc} = \tfrac{4}{9}\, \delta_a^{\,i} \delta_b^{\,j} \delta_c^{\,k} + \delta_b^{\,i} \delta_c^{\,j} \delta_a^{\,k} + \delta_c^{\,i} \delta_a^{\,j} \delta_b^{\,k} - \tfrac{2}{3} [\, \delta_b^{\,i} \delta_a^{\,j} \delta_c^{\,k} + \delta_a^{\,i} \delta_c^{\,j} \delta_b^{\,k} + \delta_c^{\,i} \delta_b^{\,j} \delta_a^{\,k} \,] \qquad (12.82)$$

and

$$f^{ijk}_{abc} = \delta_b^{\,i} \delta_c^{\,j} \delta_a^{\,k} - \delta_c^{\,i} \delta_a^{\,j} \delta_b^{\,k} \quad . \qquad (12.83)$$

(ii) Show that d^{ijk}_{abc} is symmetric under the interchange of either (i,a) and (j,b), or (i,a) and (k,c), or (j,b) and (k,c), while f^{ijk}_{abc} is antisymmetric. Both tensors satisfy the trace-less condition

$$d^{ijk}_{ibc} = f^{ijk}_{ibc} = 0 \quad .$$

(iii) From (12.48) and (12.68), one sees that $\mid p > \;=\; \mid B_3^{\,1} >$ and $< p \mid \;=\; < B_1^{\,3} \mid$. Set $j = b = 3$ and show that

$$< p \mid S_3^{\,3} \mid p > \;=\; D_s\, d^{331}_{133} + F_s\, f^{331}_{133} = \tfrac{1}{3} D_s + F_s \quad . \qquad (12.84)$$

Likewise, prove that

$$< \Lambda \mid S_3^{\,3} \mid \Lambda > \;=\; \tfrac{2}{3} D_s \quad , \quad < \Sigma \mid S_3^{\,3} \mid \Sigma > \;=\; -\tfrac{2}{3} D_s \quad ,$$
$$< \Xi \mid S_3^{\,3} \mid \Xi > \;=\; \tfrac{1}{3} D_s - F_s \quad , \qquad (12.85)$$

and therefore

$$2 < p \mid S_3^{\,3} \mid p > + 2 < \Xi \mid S_3^{\,3} \mid \Xi > \;=\; 3 < \Lambda \mid S_3^{\,3} \mid \Lambda > + < \Sigma \mid S_3^{\,3} \mid \Sigma >$$

which is the octet mass formula (12.70).

Problem 12.2. Let $J_\lambda(x)_b^{\ k}$ be either the vector or the axial-vector hadron SU_3-octet current operator. Using the notation of the previous problem, show that, depending on whether $J_\lambda(x)_b^{\ j}$ is a vector or axial-vector current,

$$\lim_{q=0} \langle B_a'^{\,i} \mid J_\lambda(x)_b^{\ j} \mid B_c^{\ k} \rangle = i(D\, d^{ijk}_{abc} + F\, f^{ijk}_{abc}) \cdot \begin{cases} u'^{\dagger} \gamma_4 \gamma_\lambda u & \text{vector} \\ u'^{\dagger} \gamma_4 \gamma_\lambda \gamma_5 u & \text{axial-vector} \end{cases} \tag{12.86}$$

where D and F are constants. Thus, if h and h' denote various hadrons, the above expression can also be written as

$$\lim_{q=0} \langle h' \mid J_\lambda(x)_b^{\ j} \mid h \rangle \equiv i \langle h' \mid \mathcal{J}_b^{\ j} \mid h \rangle \cdot \begin{cases} u'^{\dagger} \gamma_4 \gamma_\lambda u & \text{vector} \\ u'^{\dagger} \gamma_4 \gamma_\lambda \gamma_5 u & \text{axial-vector} \end{cases} . \tag{12.87}$$

Prove the following table of the reduced matrix elements $\langle h' \mid \mathcal{J}_b^{\ j} \mid h \rangle$ for $b = 2$, $j = 1$ and $b = 3$, $j = 1$. [These currents are important for the Cabibbo theory of the weak interaction.]

(The table appears on the next page.)

$h \to h'$	$\langle h' \mid \mathcal{J}_2^1 \mid h \rangle$	$\langle h' \mid \mathcal{J}_3^1 \mid h \rangle$
$n \to p$	$D - F$	0
$\Sigma^- \to \Lambda^0$	$-\sqrt{\frac{2}{3}}\, D$	0
$\Lambda^0 \to \Sigma^+$	$-\sqrt{\frac{2}{3}}\, D$	0
$\Sigma^- \to \Sigma^0$	$-\sqrt{2}\, F$	0
$\Xi^- \to \Xi^0$	$D + F$	0
$\Lambda^0 \to p$	0	$(D - 3F)/\sqrt{6}$
$\Sigma^- \to n$	0	$D + F$
$\Xi^- \to \Lambda^0$	0	$(D + 3F)/\sqrt{6}$
$\Xi^- \to \Sigma^0$	0	$(D - F)/\sqrt{2}$
$\Xi^0 \to \Sigma^+$	0	$D - F$.

Reference

Y. Ne'eman and M. Gell-Mann, eds., The Eightfold Way (New York, Benjamin, 1964).